# REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

AD-A203 122

| 1a. REPORT SECURITY CLASSIFICATION | 1b. RESTRICTIVE MARKINGS |
| --- | --- |

3. SECURITY CLASSIFICATION AUTHORITY

3. DISTRIBUTION / AVAILABILITY OF REPORT

b. DECLASSIFICATION / DOWNGRADING SCHEDULE

4. PERFORMING ORGANIZATION REPORT NUMBER(S)

5. MONITORING ORGANIZATION REPORT NUMBER(S)

AFOSR-TR- 88-1225

| 6a. NAME OF PERFORMING ORGANIZATION | 6b. OFFICE SYMBOL (If applicable) | 7a. NAME OF MONITORING ORGANIZATION |
| --- | --- | --- |
| Univ of Wisconsin | | AFOSR/NL |

| 6c. ADDRESS (City, State, and ZIP Code) | 7b. ADDRESS (City, State, and ZIP Code) |
| --- | --- |
| 1500 Highland Avenue Madison, WI 53705-2280 | Bldg 410 Bolling AFB, DC 20332-6448 |

| 8a. NAME OF FUNDING / SPONSORING ORGANIZATION | 8b. OFFICE SYMBOL (If applicable) | 9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER |
| --- | --- | --- |
| AFOSR | NL | AFOSR-87-0240 |

| 8c. ADDRESS (City, State, and ZIP Code) | 10. SOURCE OF FUNDING NUMBERS | | | |
| --- | --- | --- | --- | --- |
| Bldg. 410 Bolling AFB, DC 20332-6448 | PROGRAM ELEMENT NO. 61102F | PROJECT NO. 2313 | TASK NO A6 | WORK UNIT ACCESSION NO. |

11. TITLE (Include Security Classification)

Additivity and Auditory Pattern Analysis

12. PERSONAL AUTHOR(S)
Dr. Lutfi

| 13a. TYPE OF REPORT | 13b. TIME COVERED | 14. DATE OF REPORT (Year, Month, Day) | 15. PAGE COUNT |
| --- | --- | --- | --- |
| Annual | FROM 1 May 87 TO 30 Apr 88 | May 88 | 42 |

16. SUPPLEMENTARY NOTATION

| 17. COSATI CODES | | | 18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number) |
| --- | --- | --- | --- |
| FIELD | GROUP | SUB-GROUP | |
| | | | |
| | | | |

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

The project is designed to answer specific questions regarding listeners' ability to integrate information within and across stimulus dimensions, to extract information contained in the pattern of the acoustic signal, and to perform under conditions of stimulus uncertainty. The data are also used to determine how listeners weight the information provided by different components of the signal, and how best to package the acoustic information in frequency and/or time so that it is processed most effectively by the listener. Finally, work is undertaken to develop a computational model to summarize and predict the results of these and future experiments.

Keywords: Auditory Perception, Pattern recognition, Information Processing, Discrimination, Mathematical models. (AW)

DEC 0 9 1988

| 20. DISTRIBUTION / AVAILABILITY OF ABSTRACT | 21. ABSTRACT SECURITY CLASSIFICATION |
| --- | --- |
| ☐ UNCLASSIFIED/UNLIMITED   ☐ SAME AS RPT.   ☐ DTIC USERS | UNCLASSIFIED |
| 22a. NAME OF RESPONSIBLE INDIVIDUAL | 22b. TELEPHONE (Include Area Code) | 22c. OFFICE SYMBOL |
| Dr. John F. Tangney | (202) 767-5021 | NL |

DD Form 1473, JUN 86          Previous editions are obsolete.          SECURITY CLASSIFICATION OF THIS PAGE

UNCLASSIFIED

Annual Technical Report, 5/1/87 - 4/30/88

## Additivity and Auditory Pattern Analysis

Robert A. Lutfi, Principal Investigator

## Project Summary

Human discrimination of complex acoustic signals typically cannot be predicted from the simple sum of the discriminabilities associated with individual components of the signal. Understanding such failures of additivity is central to our understanding of complex sound perception. The goal of this project is to elucidate the rules and mechanisms whereby individual stimulus components combine to influence the detection and discrimination of complex sounds. The project is designed to answer specific questions regarding listeners' ability to *integrate* information within and across stimulus dimensions, to extract information contained in the *pattern* of the acoustic signal, and to perform under conditions of *stimulus uncertainty*. The data are also used to determine how listeners *weight* the information provided by different components of the signal, and how best to *package* the acoustic information in frequency and/or time so that it is processed most effectively by the listener. Finally, work is undertaken to develop a *computational model* to summarize and predict the results of these and future experiments.

## Statement of Work/Research Objectives

Can the perception of a complex event be reduced to the sum of its analyzable elements? This was one of the fundamental questions that occupied the minds of the earliest thinkers interested in understanding human perception. Today, of course, we are familiar with the Gestaltist's favorite illusions demonstrating that the perception of the whole is often greater than the sum of its separate parts. By demonstrating the importance of the relations among parts, the Gestalt psychologist redefined the study of perception as the study of patterns.

In contemporary psychoacoustics, the Gestaltist's influence has been made evident in pattern perception models of pitch (Goldstein, 1973; Terhardt, 1974; Wightman, 1973), localization (Searle, 1982; Perkins, Kistler and Wightman, 1986), and speech (Stevens and Blumstein, 1978). Now there is evidence that simple auditory detection, as well, frequently involves an analysis of the overall pattern of excitation produced by the signal and masker (Ahumada and Lovell, 1971; Ahumada, Marken, and Sandusky,1975; Green, 1983; Green, and Kidd, 1983; Green, and Mason, 1985; Hall, Haggard, and Fernandes, 1984; Hanna, 1984; Leek, and Watson, 1984; Lutfi, 1985, 1986; Spiegel, Picardi, and Green, 1981). The basic result of the detection studies is a failure of additivity; components of the acoustic complex affect threshold in ways that are not predicted by summing their separate effects. Failures of additivity impose severe constraints on our ability to predict the auditory system's response to complex stimuli, like speech, from the response to much simpler inputs. Thus, one of the greatest challenges confronting psychoacoustics in the years ahead is to understand the mechanisms and invariances that determine how stimulus components combine to influence auditory perception.

1

The present project adopts an approach to this problem which is both simple and direct. In all experiments, the unit of analysis is the discriminability, as measured by d', of single tone bursts that differ (on average) in level. The complex signals of these experiments are comprised of various combinations of 2 to 13 of these tone bursts distributed in frequency and/or time. On the basis of simple additivity, the discriminability of the complex is given by the vector summation rule, $d'_{complex} = (\Sigma d'^2_i)^{1/2}$, where $d'_i$ is the discriminability of the ith tone component of the complex. The vector summation rule thus provides the referent for evaluating the discriminability actually obtained. This simple approach is used to address the following specific questions regarding the processing of complex sounds:

(1) How efficiently can human observers *integrate* information within and across different stimulus dimensions?

(2) What effects do varying degrees of *stimulus uncertainty* along relevant and irrelevant dimensions have on the ability to integrate this information?

(3) How efficiently can observers extract information contained in the *pattern* of level variation across the individual components of the complex?

(4) Which components of the complex are *weighted* most heavily in the decision process?

(5) What is the best way to *package* the acoustic information in frequency and/or time so that it will be processed most effectively by the observer?

(6) What are the mechanisms underlying the discrimination of these complex sounds? Can a *computational model* be developed to account for the results?

**Research Progress**
Study 1: Magnitude Analysis
    This early experiment was designed to address two questions: How efficiently is information combined across frequency channels, and what effect does spectral uncertainty have on the ability to combine this information? The stimuli were n-tone complexes, where n ranged from 1 to 13. The frequencies of the tones were spaced at equilog intervals from 250 to 4000 Hz. Fig. 1A shows an example of one of these complexes where n is 10. In this experiment, the tones were added from low frequencies to high as n was increased (lo-pass condition). All tone complexes were played over 16-bit, audio-quality, D-to-A converters at a 20-kHz rate. The complexes were gated on and off with 5-ms, cosine-squared ramps for a total duration (from 0 voltage points) of 100 ms. On each interval of a two-interval, forced-choice trial, the individual intensities of the tones in the complex comprised a random sample of size n from one of two log-normal distributions: LOW ($M_l$ = 65 dB, $\sigma_l$ = 5 dB) and HIGH ($M_h$ = 70 dB, $\sigma_h$ = 5 dB). The value of n was fixed for each block of trials. The listener's task was to identify which interval contained the complex drawn from the HIGH distribution. Feedback was given after each trial. The recording of trial by trial data, the generation of stimuli, and all other experimental events were controlled by an IBM AT computer.
    According to the Theory of Signal Detection, optimal performance for this task as measured by d' grows as the square root of n. Specifically, $d'_{opt} = n^{1/2}\Delta/\sigma$, where $\Delta = M_h - M_l$, and $\sigma_h = \sigma_l = \sigma$. Optimal performance is the
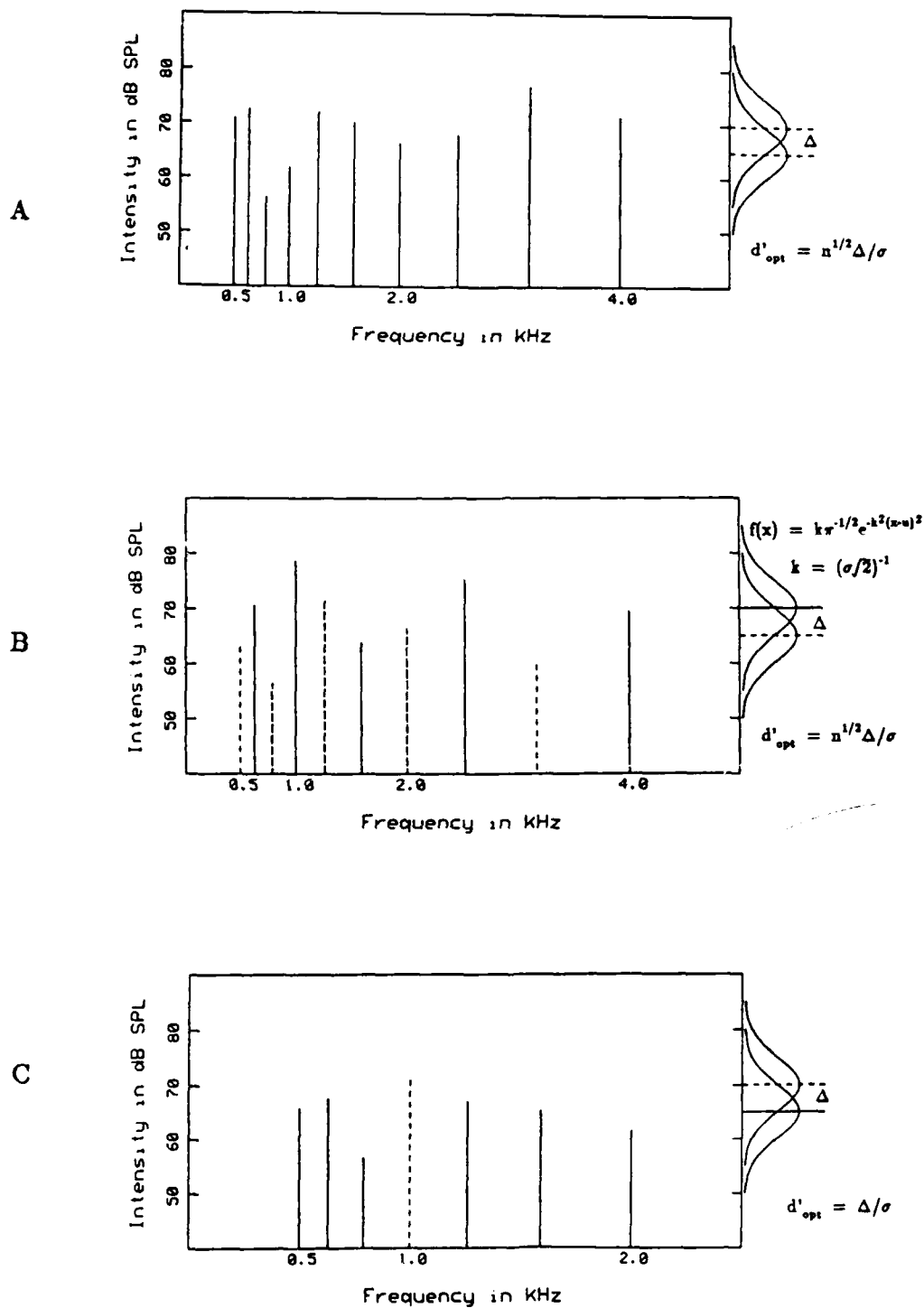
2

A

Intensity in dB SPL

80

70

60

50

Frequency in kHz

0.5  1.0  2.0  4.0

$\Delta$

$d'_{opt} = n^{1/2}\Delta/\sigma$

B

Intensity in dB SPL

80

70

60

50

Frequency in kHz

0.5  1.0  2.0  4.0

$f(x) = k\pi^{-1/2}e^{-k^2(x-u)^2}$

$k = (\sigma/2)^{-1}$

$\Delta$

$d'_{opt} = n^{1/2}\Delta/\sigma$

C

Intensity in dB SPL

80

70

60

50

Frequency in kHz

0.5  1.0  2.0

$\Delta$

$d'_{opt} = \Delta/\sigma$

FIG 1. Examples of idealized stimulus spectra used in three different experiments, (A) Magnitude analysis experiment, (B) Pattern analysis experiment, (C) Information reliability experiment. See text for further details.
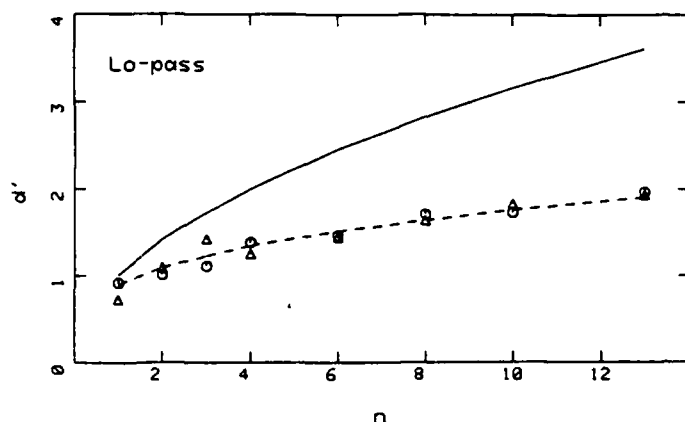
FIG 2. Integration functions from pilot experiment. Symbols represent the averages of three subjects, over 1000 trials per subject; $\sigma=5$ dB (circles), $\sigma=10$ dB (triangles). Solid line is performance of an ideal detector. Dashed line is prediction of a model. See text for further details.
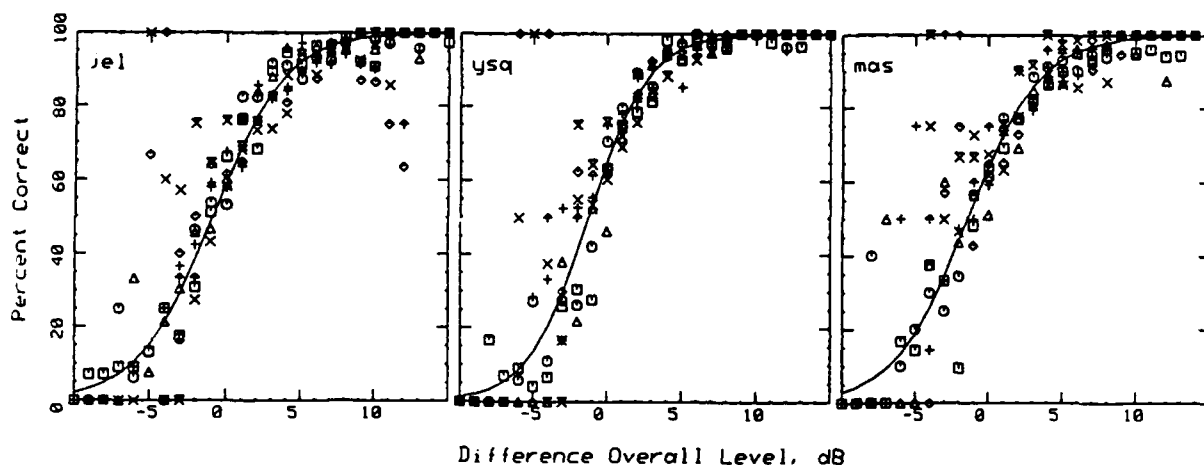


FIG 3. Psychometric functions from pilot experiment. Each panel gives data from a different subject. Plotted along the abscissa is the difference between the overall level of the two stimuli in each trial. Different symbols represent different n.
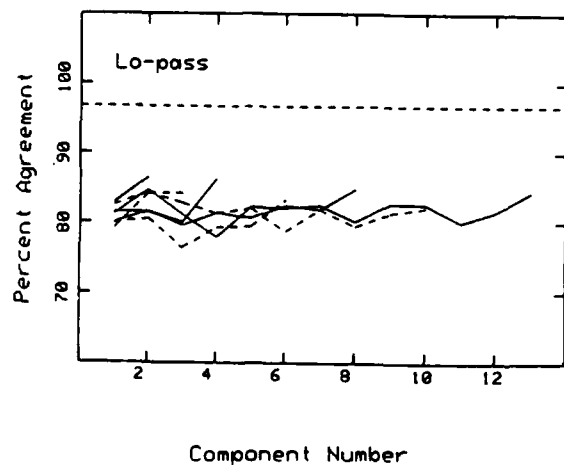


FIG 4. Weighting functions from pilot experiment. Horizontal dashed line is percent agreement for $n=1$. Other curves are percent agreement for $n>1$. Average of three subjects. See text for further details.

referent whereby we can determine how efficiently our listeners are able to make use of the information provided by the different tones of the complex. Also, the value of $\sigma$ provides an index of the degree of spectral uncertainty associated with the task - the greater the value of $\sigma$, the greater the amount of uncertainty. To measure the effects of uncertainty on listeners' ability to combine information, we simply vary $\sigma$ while being sure at the same time to adjust n or $\Delta$ so as to maintain the same level of performance from an ideal detector.

Fig. 2 shows the results of this experiment. The solid curve represents optimal performance for the task as predicted by TSD. The circles represent the average performance of three listeners, over 1000 trials per listener. The triangles represent the average performance of these same 3 listeners when both $\sigma$ and $\Delta$ were changed from 5 to 10 dB (constant $d'_{opt}$). The dashed line is the prediction of a model which will be descrbed later. For future reference, we will refer to curves drawn in these coordinates as *integration functions*. First, note that the listeners' ability to integrate information across frequencies is less than optimal. Whereas optimal performance grows at the square root of n, obtained performance grows more nearly as the cube root of n (cube root of n growth is the dashed line).[1] What is intriguing about this result is that even for small n listeners make so little use of the information available in the stimulus. Based on numerous studies of the information processing capacity of humans, we had expected that at least 7 give or take 2 components would have been processed optimally (Miller, 1963). The suboptimal performance cannot be attributed to a ceiling effect - we have since replicated the cube root of n growth in performance at a lower level of d'opt. It is also unlikely that the tones were masking one another - even with a half-octave separation between the tones performance is unchanged. Finally, the results cannot be attributed to a simple lack of training. Most of subjects are practiced musicians, some have been participating in these experiments for over a year now with little or no observable improvement in performance.

Another intriguing result is that increasing the level of stimulus uncertainty ($\sigma$ of 5 versus 10 dB) has no effect. We had expected that, in general, higher levels of uncertainty would produce poorer performance as Watson and many others have found. In fact, a $\sigma$ of 5 and 10 dB represents a fairly wide range of stimulus variability. The apparent discrepancy appears to be related to the fact that in Watson's experiments, unlike ours, there is no variation in the *difference* (level, frequency, or duration) to be discriminated; put simply, optimal performance for Watson's task is unbounded. This may be important. Most naturally occuring signals vary, thus, even an ideal detector would make errors on occasion. But this is precisely the type of stimulus variability which in our experiment has no effect. Could it be that the effects of stimulus uncertainty are largely limited to those laboratory conditions in which the difference to be discriminated is fixed, that is, in which an ideal detector makes no errors?

In this regard at least, our task appears more analogus to a traditional noise intensity discrimination task (e.g. Green, 1965) than to Watson uncertainty experiment. Our $\sigma$ could be likened to the $\sigma$ of the sampling distribution of noise energies; our $\Delta$ would be analogus to the mean difference between noise energies to be discriminated. Of course, in our experiment, $\sigma$ is generally much larger than in the noise discrimination experiment, and the form of the distributions are different as well. The important point to note however is that performance in both experiments is found to be constant for a constant $\Delta/\sigma$ ratio, reflecting a type of

---

[1] At first, this may seem inconsistent with the results of earlier studies (e.g. Green, 1960) showing square root of n growth for intensity discrimination of noise signals (where n refers to signal bandwidth). One must remember, however, that in our experiment the distributions of individual tone intensities are log-normal, thus overall intensity discrimination is a suboptimal strategy.

Weber fraction in both cases. Experiments are currently underway to partial out the relative influence of peripheral and central factors on these results.

Preliminary modelling efforts

We have now pursued several computational models to account for the results of this pilot experiment. Although these models have so far only been applied to the data of this experiment, they could in principle be applied to future results obtained in any of the experiments of this proposal. Each of these models attributes suboptimal performance to a different stage of auditory processing. The outstanding feature of these models is that, despite their differences, they all provide an equally excellent summary of the preliminary data, in each case accounting for 92% or more of the total variance. We believe that future research should be largely guided by attempts to empirically test these models. Indeed, such tests should eventually converge on a subset of stimulus conditions for which optimal performance is both predicted and obtained -- this would provide the litmus test of any model.

### Table I. Models of Information Integration

| Model | General Form | Specific Form | Growth Factor |
|---|---|---|---|
| Interchannel Correlation | $d'_n = \Delta/(\sigma^2/n+R)^{1/2}$ | $R \; \alpha \; \sigma^2$ | $[n/(1+nR/\sigma^2)]^{1/2}$ |
| Compressive Nonlinearity | $F(d'_n) = \Sigma F(d'_i)$ | $F(z) = z^p$ | $n^{1/p}$ |
| Limited Memory Capacity | $d'_n = [d'^2_1 + \Sigma (w_j d'_j)^2]^{1/2}$ | $w_j=w$ a constant | $[1+w^2(n-1)]^{1/2}$ |
| Nonoptimal Decision Strategy | -- | Difference Overall Level | approx $n^{1/3}$ |

Model 1: <u>Correlated Observations</u>. In our pilot experiment, the n elements comprising the stimulus sample are independent. The basic assumption of the correlated observations model is that the n *observations* corresponding to these elements are not independent. In effect, the model assumes that there is a source of internal (central) noise which is common to all observations. The general formulation of this model is given in Table I. Note that the general form is identical to the predictions for an ideal detector with the exception of the variance term R in the denominator. The variance term R represents the influence of the central noise in this model. In the specific form of the model, R is assumed to grow with the external variance $\sigma^2$ (i.e. the internal noise is multiplicative). The value of R providing the best fit to the data is 3.4 dB ($R^{1/2}$=1.8 dB) which is in reasonably good agreement with internal noise estimates from other types of intensity discrimination experiments (e.g. Bos and DeBoer, 1966; Durlach, 1963).

Model 2: <u>Information Compression</u>. This model allows that all observations are independent. However, it assumes all observations are subject to some nonlinear transformation both before and after they are combined. In Table I, the

4

nonlinear transformation is given by F which is assumed in this case to be a power-law. When the exponent p of the power-law is 2 there is no information loss with n and the model predicts optimal performance. For $p > 2$, there is a progressive loss of information provided by each additional observation. The result is a common form of information compression (Hafter and Dye, 1983; Lutfi, 1983; Penner, 1978, 1980; Stevens, 1936). The value of p providing the best fit to the data is 3.4. This yields a compressive exponent on n of 0.3, which again is in good agreement with values obtained in the other types of studies cited above.

Model 3: <u>Limited Memory Capacity</u>. This model emphasizes the fact that on each trial of the two-interval, forced-choice task, the observer must compare the observation made on the second interval with a memory trace of the observation made on the first interval. The trace is volatile and is assumed to deteriorate over time. In our formulation of the model (Table I), performance is optimal when the memory load is only one element (this assumes that the time between observations is small, say less than a half second). For each additional element, only a fraction $w_j$ of the information is preserved by the time the second observation interval comes along for comparison. We find excellent fits to the data when all $w_j$s are assumed equal to 0.5.

Model 4: <u>Nonoptimal Decision Strategy</u>. This model assumes no special degradation or compression of information before the decision stage. Rather, the model assumes that performance is limited by the listener's choice of a nonoptimal rule for arriving at a decision. The optimal decision strategy in our pilot experiment begins by computing a level difference between the first and second observation interval for each element of the stimulus sample. The optimal strategy is then to choose interval 1 if the sum of these differences is positive, otherwise choose interval 2. We have explored a number of alternative *non*optimal decision rules and have found one in particular that provides a very good account of the data. In this nonoptimal strategy, decisions are based simply on the *overall* level difference between the first and second observation interval (see footnote 1). This decision rule approximately yields a cube root on n growth rate as shown by the dashed line in Fig. 2.

There are several approaches that will be taken to test among these models. Many tests will simply involve the manipulation of variables explicitly or implicitly defined in the mathematical formulations of the models. These variables include the variance of the distribution of members within each stimulus category (holding $d'_{opt}$ constant), the number of members, the number of categories, the mathematical form of the distributions, and the size of the sample randomly drawn on each trial. Other tests will involve various manipulations in stimulus parameters and various ways of "packaging" the information presented to observers. For instance, the tones from signal and nonsignal distributions will be intermingled in frequency and time in various ways to form different classes of spectral-temporal patterns (Study 2). The final approach will be to evaluate the models based on trial-by-trial analyses of the listeners' responses. This latter approach is discussed in greater detail below.

Trial-by-trial analyses

Each of the models we have described makes a specific prediction regarding the mathematical form of the integration function. Unfortunately, the differences among these functions are so small that they cannot be resolved within the measurement error of our experiment. In this situation, we resort to analyzing the models' predictions for the trial-by-trial data.

Consider for example the predictions of Model 4. According to this model, the listener responds to the interval perceived to have the higher overall level.

Thus, on trials in which the HIGH sample has the higher overall level the listener will usually respond correctly, on trials in which the LOW sample has the higher overall level, the listener should more often respond incorrectly. Indeed, if overall level is the cue used by listeners, then the trial-by-trial data across all conditions of the experiment should converge on a single *psychometric function*; the abscissa for this function would be the difference between the overall level of HIGH and LOW samples on each trial. Fig 3. shows the results this analysis. To obtain a percent correct value at each level difference, the trial-by-trial data were accumulated into 1 dB bins; thus the percent correct at say 5 dB is actually the percent correct for all trials in which the level difference between the HIGH and LOW samples was between 4.5 and 5.5 dB. Each panel represents the data from a different subject; the different symbols correspond to the different sample sizes (n). The solid curve in each panel is the best fitting logistic (see Bush, 1963). The data do indeed tend to converge on a single psychometric function. Note also that for a performance level of 75% correct, the overall level difference for all subjects is near 1 dB - a normal difference limen for intensity in the 2IFC procedure. This analysis provides a necessary, not a sufficient, test of model 4. It demonstrates nonetheless how the trial-by-trial data may be used to gain additional insights into the processes underlying discrimination performance.

Another use of the trial-by-trial data is to provide stimulus *weighting functions*. These functions are intended to specify the relative contribution of different stimulus elements to the decision process (see question 3). The method we have chosen is simply to count the agreements between the response on each trial and the level difference of the ith element on each trial. For example, if on a given trial the level of the ith element is higher on the second interval, and if the response is to the second interval, then the response is scored as an agreement. Fig. 4 shows the percent agreements for each of the elements as derived from the trial-by-trial data of the pilot experiment. Only those trials in which the ith level difference exceeded 5 dB were included in this analysis. This restriction was implemented to eliminate possible disagreements resulting from the listeners inability to discriminate the level difference. Now suppose that when all thirteen elements are played (n=13), the listener attends exclusively to the thirteenth element. The percent agreements in this case should equal the percent agreements when only one element was played (n=1). The percent agreements will be less than this to the extent that the listener attends to the other n-1 components. The horizontal dashed line gives the percent agreements for n=1. The results give little evidence that listeners differentially weight the various frequencies that comprise these signals. This is not too surprising since all elements constitute equally reliable sources of information. We would not expect this to be true when different reliabilites are associated with each element (study 5).

Study 2. Pattern analysis

The relevant information distinguishing many naturally occuring signals is contained not only in overall intensity differences across signals, but also in the *pattern* of intensity variations across frequency within each signal. The next study focused on listeners' ability to perform spectral pattern analysis. All conditions were identical to the pilot experiment described earlier except on each interval of the 2IFC trial, half the tones were drawn from the HIGH distribution and the other half were drawn from the LOW distribution. On one interval, the odd numbered tones were drawn from the HIGH distribution, the even numbered tones from the LOW (See Fig. 1B). On the other interval, the reverse was true. The listener's task was to identify the interval in which the odd numbered tones were drawn from the HIGH distribution. To insure that the subject's response were
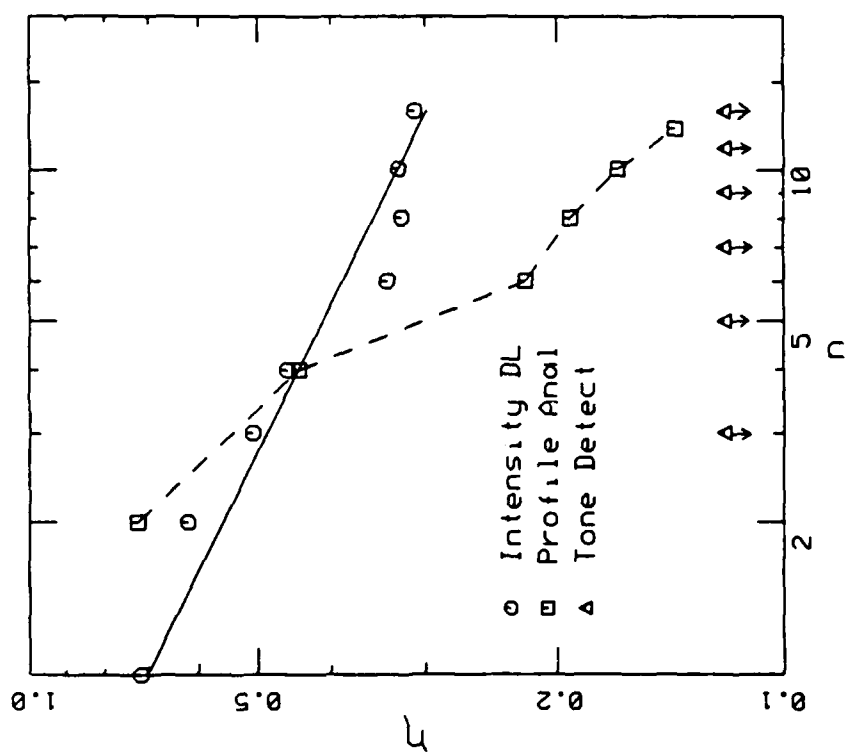
6

Fig. 5. Discrimination efficiency as a function of number of components in three experiments.

based on spectral pattern analysis, we also roved the overall level of the stimulus on each presentation (see Green, 1988).

According to the Theory of Signal Detection, optimal performance for this task is identical to that for the earlier magnitude discrimination task. -- The performance of the ideal detector is unaffected by how the information is packaged in frequency and/or time. It is of interest, therefore, to compare the performance of our subjects in this task to their performance in the earlier magnitude discrimination task. Both the trace memory model and the correlated observations model predict that human performance in the pattern analysis task will be near optimal. There is no decay of memory because the relevent comparisons are between the intensities of components that occur simultaneously with one another. There is no effect of common internal noise because the common noise is subtracted out in the differencing of components. In contrast, the nonoptimal decision model predicts that human pattern analysis will be worse than magnitude analysis.

The results of this experiment (squares) are compared with the results of the magnitude analysis experiment (circles) in Fig. 5. The data are plotted as measures of discrimination efficiency, eta $= (d'_{obt}/d'_{opt})^2$. It is clear that spectral pattern analysis is significantly poorer for our subjects than simple intensity discrimination. The difference increases with the number of components in the stimulus. Though the results support the nonoptimal decision model, we can not rule out the possibility that roving overall stimulus level may have had a detrimental effect on performance. Experiments are underway to test this hypothesis. Additional tests of these models will involve discrimination of patterns across both frequency and time.

## Study 3. Differential Weighting of Frequency Components

Not all frequencies comprising naturally occuring sounds be expected to carry the same amount of information for discrimination. Obviously, those frequency components conveying the greatest amount of information should be given greatest weight in the decision process. Study 3 investigated listeners' ability to select weights appropriate to the information content of the individual frequency components of the complex. The first experiment represented an extreme case in which all information relevant to classification was contained in a single component, the one at 1 kHz. On one interval of the 2IFC trial, the level of the 1-kHz component was drawn from the HIGH distribution. On the other interval, the level of this component was drawn from the LOW distribution. The levels of all other tones on both intervals was drawn from the LOW distribution (See Fig. 1C). The listener's task was to select the interval in which the 1-kHz component was selected from the HIGH distribution. We were quite surprised to find that several of our best subjects could not perform above chance on this task (triangles of Fig. 5), even after considerable practice. We had expected that subjects' performance would be near optimal. They would only need to focus their attention on the critical band containing the single 1-kHz component and ignore all other components. Apparently they were unable to do this. This represents a rather severe departure from the critical band principle, one that may be related to recent results obtained by Neff and Green (1987). We intend to pursue this result further by increasing the information (the $d'_{opt}$) provided by the 1-kHz component, and by slowly increasing the information conveyed by the other components. Weighting functions derived from the trial-by-trial data should indicate whether or not listeners selectively weight the individual frequency components according to their information content.

7

## Publications, Abstracts and Talks supported by AFOSR

Scientific Journal Articles

Lutfi, R.A. (1988a). "Interpreting measures of frequency selectivity: Is forward masking special?," **J. Acoust. Soc. Am. 83**, 163-177.

Lutfi, R.A. (1988b). "Complex interactions between pairs of forward maskers," **Hearing Research**, in press.

Abstracts and Talks

Lutfi, R.A. (1987). "Additivity and auditory pattern analysis," Presentation given at the AFOSR Program in Auditory Pattern Recognition. December meeting.

Lutfi, R.A. (1988c). "Discrimination of random spectra by normal-hearing listeners," in **Abstracts of the 11th Midwinter Research Meeting, ARO.** Edited by D.J. Lim, p. 209.

Lutfi, R.A. (1988d). "A detection-theoretic approach to complex sound perception," Presentation given at the Psychophysics Laboratory at the University of Florida, Gainesville.

Papers in Preparation

Lutfi, R.A. (1988e). "Informational processing of complex sound: I. Magnitude analysis," J. Acoust. Soc. Am. or J. Exp. Psychol., Human Percept & Perf.

Bibliography

Ahumada, A. Jr., and Lovell, J. (1971). "Stimulus features in signal detection," J. Acoust. Soc. Am. 49, 1751-1756.

Ahumada, A., Jr., Marken, R., and Sandusky, A. (1975). "Time and frequency analysis of auditory signal detection," J. Acoust. Soc. Am. 57, 385-390.

Berg, B.G. (1987). "Internal noise in auditory decision tasks," Doctoral Dissertation, Indiana University.

Bos, C.E., and de Boer, E. (1966). "Masking and discrimination," J. Acoust. Soc. Am. 39, 708-715.

Bush, R.R. (1963). "Estimation and evaluation," in Handbook of Mathematical Psychology, edited by R.D. Luce, R.R. Bush, and E. Galanter, (John Wiley and Sons, New York), pp. 429-469.

Durlach, N.I. (1963). "Equalization and cancellation theory of binaural masking-level differences," J. Acoust. Soc. Am. 35,1206-1218.

Durlach, N.I., Braida, L.D., and Ito, Y. (1986). "Towards a model for discrimination of broadband signals," J. Acoust. Soc. Am. 80, 63-72.

Goldstein, J.L. (1973). "An optimum processor theory for the central formation of pitch of complex tones," J. Acoust. Soc. Am. 54, 1496-1516.

Green, D.M. (1958). "Detection of multiple component signals in noise," J. Acoust. Soc. Am. 30, 904-911.

Green, D.M. (1960). "Auditory detection of a noise signal," J. Acoust. Soc. Am. 32, 121-131.

Green, D.M. (1964). "Consistency of auditory judgments," Psych. Rev. 71, 392-407.

Green, D.M. (1983). "Profile analysis: A different view of auditory intensity discrimination," Amer. Psychol. 38, 133-142.

Green, D.M. (1988). Profile Analysis: Auditory Intensity Discrimination, (Oxford University Press, New York).

Green, D.M. and Kidd, G. Jr. (1983). "Further studies of auditory profile analysis," J. Acoust. Soc. Am. 73, 1260-1265.

Green, D.M. and Mason, C.R. (1985). "Auditory profile analysis: Frequency, phase, and Weber's Law," J. Acoust. Soc. Am. 77, 1155-1161.

Green, D.M., and Swets, J.A. (1966). Signal Detection Theory and Psychophysics (Krieger: Huntington, New York).

Hall, J.W., Haggard, M.P. and Fernandes, M.A. (1984). "Detection in noise by spectro-temporal analysis," J. Acoust. Soc. Am. 76, 50-56.

Hanna, T.E. (1984). "Discrimination of reproducible noise as a function of bandwidth and duration," Percept. Psychophys. 36, 409-416.

Leek, M.R. and Watson, C.S. (1984). "Learning to detect auditory pattern components," J. Acoust. Soc. Am. 76, 1037-1044.

Lutfi, R.A. (1983). "Additivity of simultaneous masking," J. Acoust. Soc. Am. 73, 262-267.

Lutfi, R.A. (1985). "A power-law transformation predicting masking by sounds with complex spectra," J. Acoust. Soc. Am. 77, 2128-2135.

Lutfi, R.A. (1985). "A power-law transformation predicting masking by sounds with complex spectra," J. Acoust. Soc. Am. 77, 2128-2135.

Lutfi, R.A. (1986). "Two- versus four-tone masking, revisited," J. Acoust. Soc. Am., 80 422-428.

Lutfi, R.A. (1988). "Discrimination of random spectra by normal-hearing listeners," in Abstracts of the 11th Midwinter Research Meeting, ARO, Edited by D.J. Lim, p. 209.

Miller, G.A. (1963). "The magical number seven, plus or minus two: Some limits on our capacity for processing information," in Readings in Mathematical Psychology, edited by R.D. Luce, R.R. Bush, and E. Galanter, (John Wiley and Sons, New York), pp 135-151.

Neff, D.L. and Green, D.M. (19?7). "Masking produced by spectral uncertainty with multicomponent maskers," Perception & Psychophysics 41, 409-415.

Penner, M.J. (1978). "A power-law transformation resulting in a class of short-term integrators that produce time-intensity trades for noise bursts," J. Acoust. Soc. Am. 63, 195-201.

Penner, M.J. (1980). "The coding of intensity and the interaction of forward and backward masking," J. Acoust. Soc. Am. 67, 608-616.

Penner, M.J. and Shiffrin, R.M. (1980). "Nonlinearities in the coding of intensity within the context of a temporal summation model," J. Acoust. Soc. Am. 67, 617-627.

Perkins, M.E., Kistler, D.J. and Wightman, F.L. (1986). "Spectral pattern recognition and sound localization," J. Acoust. Soc. Am. submitted.

Robinson, D.E., and Berg, B.G. (1986). "A partitioned variance model for multiple observations," Paper presented at the 19th Annual Meeting of the Society for Mathematical Psychology, Cambridge, MA.

Searle, C.L. (1982). "A model of auditory localization: Peripheral constraints," In Localization of Sound: Theory and Applications, Edited by R.W. Gatehouse (Amphora Press: Groton, CT).

Spiegel, M.F., Picardi, M.C. and Green, D.M. (1981). "Signal and masker uncertainty in intensity discrimination," J. Acoust. Soc. Am. 70, 1015-1019.

Stevens, K.N. and Blumstein, S.E. (1978). "Invariant cues for place of articulation in stop consonants," J. Acoust. Soc. Am. 64, 1358-1368.

Stevens, S.S. (1936). "A scale for the measurement of psychological magnitude: loudness," Psychol. Rev. 43, 405-416.

Terhardt, E. (1974). "Pitch, consonance and harmony," J. Acoust. Soc. Am. 55, 1061-1069.

Watson, C.S. and Kelly, W.J. (1981). "The role of stimulus uncertainty in the discrimination of auditory patterns," in Auditory and Visual Pattern Recognition, edited by D.J. Getty and J.H. Howard, Jr., (Lawrence Erlbaum and Associates, Hillsdale).

Watson, C.S., Kelly, W.J. and Wroton, H.W. (1976). "Factors in the discrimination of tonal patterns II: Selective attention and learning under various levels of stimulus uncertainty," J. Acoust. Soc Am. 60, 1176-1186.

Wightman, F.L. (1973). "The pattern-transformation model of pitch," J. Acoust. Soc. Am. 54, 407-416.

# Interpreting measures of frequency selectivity: Is forward masking special?[a]

Robert A. Lutfi

*Waisman Center, University of Wisconsin, Madison, Wisconsin 53705*

In a previous article [Lutfi, J. Acoust. Soc. Am. **76**, 1045–1050 (1984)], the following relation was used to predict measures of frequency selectivity obtained in forward masking from measures obtained in simultaneous masking: $F(g) = G + H(g) - H(0)$, where, for a given masker level, $F$ is the amount of forward masking (in dB) as a function of signal-masker frequency separation $(g)$, $H$ is the amount of simultaneous masking, and $G$ is the amount of forward masking for $g = 0$. In the present study, the relation was tested for a wider range of signal and masker frequencies, masker levels, and signal delays. The relation described thresholds from all conditions well with the inclusion of one free parameter $\lambda$ corresponding to a constant frequency increment, $F(g) = G + H(g + \lambda) - H(\lambda)$. The parameter $\lambda$ was required to account for observed shifts in the frequency of maximum forward masking. It is argued that a single tuning mechanism can account for commonly observed differences between simultaneous- and forward-masked measures of frequency selectivity.

PACS numbers: 43.66.Ba, 43.66.Dc, 43.66.Fe [WAY]

## INTRODUCTION

Forward masking refers to the elevation in the threshold of a signal presented shortly after the masker has terminated. The residual effect of the masker offers a means of measuring auditory frequency selectivity free from intrusive interactions that may occur when the signal and the masker are played simultaneously (e.g., Egan and Hake, 1950; Greenwood, 1971). Unfortunately, differences exist among simultaneous- and forward-masked measures that, after many studies, are still not well understood. By far, the largest differences, and those that have received the most attention, are observed among psychophysical tuning curves. This estimate of frequency selectivity gives the level of the masker at each frequency necessary to mask a fixed-level, fixed-frequency signal. Typically, tuning curves are observed to be narrower when measured in forward masking than when measured in simultaneous masking (Duifhuis, 1976; Houtgast, 1972, 1974; Lutfi, 1984; Moore *et al.*, 1984; Moore, 1978; Weber, 1983; Wightman *et al.*, 1977). There is little agreement regarding the mechanisms underlying this result, although it has commonly been assumed that at least two separate, frequency-selective processes are involved (Duifhuis, 1976; Houtgast, 1972, 1974; Moore, 1978; O'Loughlin and Moore, 1981; Terry and Moore, 1977; Weber, 1983; Wightman *et al.*, 1977). Forward masking is believed to be fundamentally different from simultaneous masking in that it reflects the operation of these additional frequency selective processes.

The decision to invoke additional tuning mechanisms came after physiological studies had accumulated evidence of suppression from single-unit recordings in the cat's audi-

tory nerve (Kiang *et al.*, 1965), and evidence for a physiologically vulnerable "second filter" (Evans, 1975). Apparent similarities suggested possible connections between these physiological observations and the differences observed among psychophysical tuning curves. Weber (1983) has reviewed three such theories in detail and has rejected one of them. Later interpretations were to implicate "off-frequency listening" (O'Loughlin and Moore, 1981) and "cuing" effects (Terry and Moore, 1977; Moore, 1978). However, the frequency-dependent nature of these effects preserved the general assumption that differences among tuning curves somehow reflect additional frequency-selective processes operating in forward masking.

More recently, articles have begun to question the extent to which additional tuning mechanisms are involved. In a contemporary review of the literature of frequency selectivity, Jesteadt and Norton (1985) note that forward-masking tuning curves may broaden markedly at high signal levels, while simultaneous-masking tuning curves appear to change little (Stelmachowicz and Jesteadt, 1984). They suggest that forward-masking tuning curves may be narrower than simultaneous-masking tuning curves only for moderate- and low-level signals; for high-level signals, forward-masking tuning curves might actually be broader. Subsequent data of Moore and Glasberg (1986) indicate that the difference between simultaneous- and forward-masking tuning curves is reduced slightly at high signal levels. Nelson and Freyman (1984) report a similar, perhaps related, broadening of tuning curves with increasing signal delay (also see Kidd and Feth, 1981; Small and Busse, 1980). They show that, if signal level is selected to equate the tips of the tuning curves, the tuning curves do not change significantly with signal delay. Bacon and Moore (1986) found that the difference between simultaneous- and forward-masking tuning curves also depends on the temporal placement of the signal within the simultaneous masker. When

the signal occurs at either end of the simultaneous masker (the trailing end being the typical placement in tuning curve experiments), forward-masking tuning curves do appear significantly narrower. However, when the signal occurs in the temporal center of the simultaneous masker, this difference is much reduced. Even when tuning curves have shown large differences in simultaneous and forward masking, the role of additional tuning mechanisms has been questioned since other measures of frequency selectivity fail to show such large differences (Lutfi, 1984; Weber and Lutfi, 1986). One such measure, the filter function, gives signal threshold as a function of masker frequency for a fixed-level masker. Lutfi (1984) reports filter functions that are essentially parallel in simultaneous and forward masking over a wide range of masker levels.

The recent studies complicate the interpretation of forward-masking tuning curves. They suggest that differences in measures, once thought to reflect the operation of additional tuning mechanisms, may, in large part, be attributed to interactions among the effects of masker frequency, masker level, and signal delay that are peculiar to the tuning curve experiment. Presently, it is difficult to determine the influence of such interactions. The previous studies have generally focused on the effects of one or two of these factors while holding the remaining factor(s) constant. Specific values chosen for the remaining factors may, therefore, have played a role in producing the observed differences among tuning curves. The purpose of the present study is to investigate the interaction among all three factors rather than to provide a fine-grain analysis of any one. Comparatively few masker frequencies were used so that tuning curves and filter functions could be obtained for a wider range of signal frequencies, masker levels, and signal delays than has been typical of any one study. Based on the results, it is argued that differences commonly observed between simultaneous- and forward-masking tuning curves are largely epiphenomena of the tuning curve procedure.

## I. METHOD

Filter functions and tuning curves were obtained in simultaneous and forward masking by measuring threshold for a brief sinusoidal signal in the presence of a variable-frequency, narrow-band noise masker. Filter functions were obtained for three signal frequencies ($f_s = 0.5$, 1.0, and 2.0 kHz); tuning curves were obtained for two signal frequencies ($f_s = 0.5$ and 2.0 kHz). The masker frequencies ($f$) varied in proportion to the signal frequency $(f - f_s)/f_s = -0.3, -0.2, -0.05, 0.0, 0.05, 0.1,$ and $0.2$. In simultaneous masking, the offset of the signal (0 voltage point) coincided with the offset of the masker. In forward masking, the *onset* of the signal followed the offset of the masker by 5, 10, 20, or 40 ms. Masker level varied from 30–90 dB SPL depending on the particular combination of masker frequency and signal delay. Complete filter functions were obtained for masker levels of 50–80 dB SPL. These data were then used to derive tuning curves at signal levels of 30–60 dB SPL. The details of this derivation are described in Sec. II. Not all filter shapes and tuning curves were obtained for all possible combinations of level and signal delay.

In simultaneous masking, for relative masker frequencies $(f - f_s)/f_s = -0.3$ and $-0.2$, a control measure was taken to prevent the detection of aural combination bands generated by signal–masker interaction (see Greenwood, 1971). A low-level band of noise (50 Hz wide and 30 dB below the level of the primary masker) was gated on and off in the same manner as the signal. The center frequency of the additional noise band was set equal to the center frequency of the most audible aural combination band at $2f - f_s$ (Greenwood, 1971). The amount of masking produced by this additional noise band alone was always 25 dB or more below that produced by the primary masker, and so it was not expected to produce any additional masking of the signal.

### A. Stimuli

The signal was a 10-ms sinusoid, shaped with 5-ms, Kaiser ($W_a = 0.2$) onset and offset ramps (Childers and Durling, 1975). This ramp has the desirable property that the spectral sidelobes are more than 70 dB down from the primary lobe within 20% of the primary lobe center frequency. The narrow-band noise maskers had 3-dB bandwidths of 50 Hz. They were gated on and off with 5-ms Kaiser ramps for a total duration (between the 0 voltage points) of 200 ms. All stimuli were digitally (PDP-11/40) synthesized and output through 14-bit DACs, low-pass filtered at the 4-kHz cutoff frequency of Unigon (model LP-120, 120 dB/oct) and Khron-hite (model 3343, 96 dB/oct) filters. The narrow-band noise maskers were randomly sampled from a 3-s noise file. The levels of all stimuli were controlled by programmable attenuators, and all stimuli were presented over TDH-49 headphones (with 65001 cushions) to the right ear of subjects seated in a IAC, double-wall, sound-attenuated chamber.

### B. Procedure

In all conditions, signal threshold was the dependent variable. Signal thresholds were obtained in daily 2-h sessions using a two-interval, forced-choice, adaptive procedure (see Levitt, 1971). Threshold estimates were based on the average of the last eight reversals in each adaptive run after the first two reversals had been rejected. Five such estimates were obtained on different days for each condition of the experiment. The lowest and highest of the five estimates were rejected and the remaining three were averaged to arrive at the final threshold estimate. The standard error of the trimmed mean exceeded 3 dB for 5% of the cases (see Barnett and Lewis, 1978 for information regarding the use of trimmed means). For all subjects, the pattern of results was quite similar. Therefore, the data were further averaged across subjects.

Four normal-hearing individuals were paid observers in each phase of the study, although the same four observers did not participate in each phase. One subject was unable to continue after data had been collected for the 2.0-kHz signal. Data for the 0.5-kHz signal were, therefore, collected with a replacement subject. A second replacement was required before collecting data for the 1.0-kHz signal. The ages of the

subjects ranged from 21–30 years. Each subject was tested individually.

## II. RESULTS

### A. Filter functions in simultaneous and forward masking as a function of signal frequency

Figure 1 shows simultaneous- and forward-masked thresholds as a function of the relative masker frequency for each of the three signal frequencies. Masker level is 80 dB SPL and signal delay in forward masking is 5 ms. Delaying the signal results in an overall reduction in masked threshold, but the reduction in threshold does not always appear to be the same at each masker frequency. For instance, threshold for the 2.0-kHz signal is greatest in simultaneous masking when the masker frequency is 2.0 kHz. In forward masking, however, the 1.9-kHz masker produces the highest threshold. The effect is also evident for the 0.5- and 1.0-kHz signals. In each case, the maximum masking frequency (MMF) in forward masking occurs at a frequency just below that obtained in simultaneous masking. Similar shifts in the MMF in forward masking have been reported previously by a number of investigators (Ehmer and Ehmer, 1969; Kidd and Feth, 1981; Munson and Gardner, 1950; Nelson and Freyman, 1984; Vogten, 1978a,b; Widin and Viemeister, 1979a,b; Zwicker and Jaroszewski, 1982). These shifts have been examined most extensively in the context of psychophysical tuning curve experiments. Therefore, discussion of them is reserved for the section on tuning curves.

The simultaneous-masked thresholds are described adequately, on the selected coordinates, by filter functions with two linear segments. Expressing relative masker frequency as $g = (f - f_s)/f_s$, the filter functions are of the form

$$H(g) = \begin{cases} T_{max} - \beta_l |g - \alpha|, & g < \alpha, \\ T_{max} - \beta_u |g - \alpha|, & g > \alpha, \end{cases} \quad (1)$$

where $\beta_u$ and $\beta_l$ are, respectively, the unsigned slopes of the upper and lower branches of the function, $\alpha$ corresponds to the frequency at the break point, and $T_{max}$ is signal threshold (in dB) at the breakpoint. The parameter $\alpha$ allows the MMF to be estimated slightly above or below the signal frequency. The curves drawn through the simultaneous-masked thresholds were obtained by selecting values of $\beta_u$, $\beta_l$, $\alpha$, and $T_{max}$ satisfying the least-squares criterion. The results of the regression for the individual and mean data are shown in Table I (80-dB masker level). Each curve represents the regression of four parameters on only seven points, so the proportion of variance accounted for ($r^2$) is predictably high. In subsequent analysis, these simultaneous-masking filter functions will be used to predict the forward-masked thresholds.

The degree of frequency selectivity exhibited by simultaneous-masking filter functions is estimated by the steepness of the unsigned slopes, $\beta_u$ and $\beta_l$. For the 1.0- and 2.0-kHz signals, the low-frequency slope is small relative to the high-frequency slope reflecting the familiar upward spread of masking. The average 3-dB bandwidths derived from the slopes are 83, 83, and 252 Hz, respectively, for the 0.5-, 1.0-, and 2.0-kHz signals. These values are in reasonable agree-
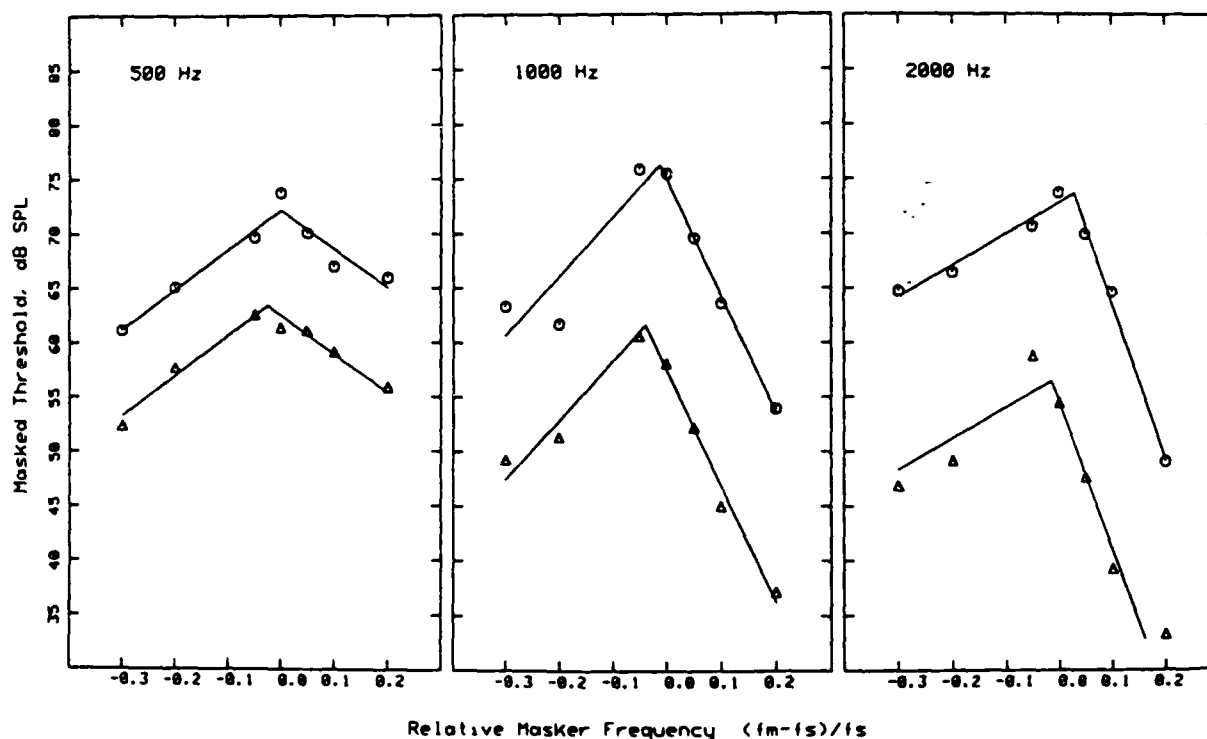


FIG. 1. Simultaneous- and forward-masked thresholds (circles and triangles, respectively) as a function of relative masker frequency $g$ for three signal frequencies $f_s = 500$, 1000, and 2000 Hz. Signal delay is 5 ms. The data are the averaged threshold of four subjects. The filter functions drawn through the simultaneous-masked thresholds were obtained by least-squares regression according to Eq. (1). The forward-masking filter functions were derived from the simultaneous-masking filter functions according to Eq. (2).

| Subject | $f_i$ Hz | Masker level dB SPL | Parameters $\alpha$ | $\beta_u$ dB | $\beta_l$ dB | $T_{max}$ dB SPL | $r^2$ | rms error dB |
|---|---|---|---|---|---|---|---|---|
| JL | 500 | 80 | 0.025 | 68.8 | 25.3 | 76.1 | 0.892 | 1.2 |
|  |  | 70 | 0.050 | 55.3 | 27.0 | 66.9 | 0.920 | 1.1 |
|  |  | 60 | 0.067 | 56.7 | 56.1 | 60.3 | 0.934 | 1.9 |
|  |  | 50 | 0.021 | 26.2 | 37.1 | 49.2 | 0.825 | 1.2 |
|  | 1000 | 80 | −0.033 | 72.9 | 64.2 | 74.5 | 0.732 | 3.7 |
|  | 2000 | 80 | 0.021 | 131.0 | 29.4 | 73.6 | 0.995 | 0.5 |
|  |  | 70 | 0.037 | 101.0 | 44.8 | 62.2 | 0.914 | 1.7 |
|  |  | 60 | 0.074 | 146.0 | 59.3 | 56.3 | 0.979 | 1.4 |
|  |  | 50 | 0.050 | 73.8 | 42.3 | 46.1 | 0.946 | 0.9 |
| DO | 500 | 80 | 0.000 | 30.7 | 29.9 | 70.5 | 0.881 | 1.1 |
|  |  | 70 | 0.056 | 65.8 | 35.3 | 62.6 | 0.842 | 1.8 |
|  |  | 60 | 0.062 | 35.0 | 34.7 | 53.2 | 0.736 | 2.5 |
|  |  | 50 | 0.049 | 2.7 | 35.1 | 42.7 | 0.891 | 0.9 |
|  | 1000 | 80 | 0.002 | 94.8 | 55.8 | 81.3 | 0.892 | 2.1 |
|  | 2000 | 80 | 0.091 | 273.0 | 51.5 | 76.8 | 0.927 | 2.6 |
|  |  | 70 | 0.042 | 197.0 | 94.5 | 70.0 | 0.978 | 2.0 |
|  |  | 60 | 0.046 | 143.0 | 105.0 | 56.1 | 0.902 | 3.5 |
|  |  | 50 | 0.049 | 137.0 | 109.0 | 50.4 | 0.972 | 1.2 |
| BL | 500 | 80 | −0.034 | 31.7 | 62.3 | 73.8 | 0.962 | 1.0 |
|  |  | 70 | 0.050 | 47.2 | 38.0 | 62.9 | 0.977 | 0.7 |
|  |  | 60 | 0.029 | 9.3 | 51.7 | 53.0 | 0.982 | 0.7 |
|  |  | 50 | 0.082 | 50.0 | 53.9 | 47.6 | 0.988 | 0.9 |
|  | 1000 | 80 | 0.013 | 133.5 | 50.8 | 77.0 | 0.873 | 2.8 |
| FB | 2000 | 80 | 0.000 | 126.0 | 27.6 | 73.7 | 0.975 | 1.3 |
|  |  | 70 | 0.000 | 115.0 | 47.5 | 63.1 | 0.967 | 1.1 |
|  |  | 60 | 0.000 | 80.0 | 67.2 | 54.4 | 0.874 | 2.5 |
|  |  | 50 | 0.017 | 87.1 | 66.2 | 47.2 | 0.944 | 1.3 |
| DD | 500 | 80 | 0.000 | 22.8 | 46.3 | 69.7 | 0.829 | 2.2 |
|  |  | 70 | 0.000 | 53.4 | 50.8 | 60.8 | 0.951 | 0.9 |
|  |  | 60 | 0.030 | 4.4 | 27.4 | 49.0 | 0.922 | 1.1 |
|  |  | 50 | 0.000 | 11.4 | 164.0 | 42.5 | 0.997 | 0.7 |
| BZ | 1000 | 80 | −0.013 | 142.9 | 39.2 | 76.6 | 0.979 | 1.4 |
| DD | 2000 | 80 | 0.049 | 145.0 | 15.4 | 71.2 | 0.958 | 1.5 |
|  |  | 70 | 0.034 | 144.0 | 46.4 | 66.1 | 0.959 | 1.5 |
|  |  | 60 | 0.021 | 105.0 | 69.0 | 54.2 | 0.988 | 0.9 |
|  |  | 50 | −0.005 | 54.7 | 71.6 | 43.5 | 0.851 | 2.1 |
| Mean | 500 | 80 | 0.000 | 35.7 | 36.9 | 72.3 | 0.993 | 0.5 |
|  |  | 70 | 0.040 | 57.1 | 37.6 | 63.4 | 0.963 | 1.0 |
|  |  | 60 | 0.050 | 26.9 | 41.5 | 53.9 | 0.982 | 0.7 |
|  |  | 50 | 0.042 | 14.1 | 43.0 | 45.0 | 0.951 | 0.9 |
|  | 1000 | 80 | −0.012 | 107.0 | 54.5 | 76.4 | 0.921 | 2.1 |
|  | 2000 | 80 | 0.030 | 143.0 | 28.5 | 73.7 | 0.990 | 1.0 |
|  |  | 70 | 0.030 | 138.0 | 58.4 | 65.3 | 0.991 | 1.1 |
|  |  | 60 | 0.035 | 111.0 | 72.0 | 55.0 | 0.961 | 1.5 |
|  |  | 50 | 0.032 | 83.1 | 74.8 | 46.9 | 0.954 | 1.4 |

ment with bandwidth estimates of the auditory filter obtained in simultaneous masking with less intense maskers[1] Weber (1977), for instance, reports 3-dB bandwidths of 97 and 217 Hz, respectively, for a 1.0- and 2.0-kHz signal. Patterson (1976) obtained a 3-dB bandwidth of 69 Hz for a 0.5-kHz signal.

We wish to determine whether or not the degree of frequency selectivity, as indicated by the slopes of the filter functions, differs significantly in forward masking. If the filter function in simultaneous masking is designated $H(g)$, then all filter functions having identical slopes are of the form

$$F(g) = H(g + \lambda) + \zeta, \tag{2}$$

where $\lambda$ and $\zeta$ are constants. The parameter $\lambda$ merely represents a shift in the breakpoint frequency; it provides an estimate of the shift in the MMF in forward masking. The parameter $\zeta$ gives the corresponding change in $T_{max}$. The simultaneous-masking filter functions were used to predict the forward-masked thresholds by estimating the constants $\lambda$ and $\zeta$ in Eq. (2). The results are the curves drawn through the forward-masked thresholds in Fig. 1. Table II ($M = 80$ dB, $t = 5$ ms) gives, for the individual and mean data, the values of $\lambda$ and $\zeta$ satisfying the least-squares criterion. For

TABLE II. Parameters $\lambda$ and $\zeta$ (entries) for forward-masking filter functions. Note that the $r$ value in the right-hand column refers to the proportion of total variability accounted for in all conditions of the corresponding row.

| Subject | $f_i$ Hz | Masker level | | | | | | | | $r$ | rms error dB |
| | | 50 | | 60 | | 70 | | 80 dB SPL | | | |
| | | $\lambda$ | $\zeta$ | $\lambda$ | $\zeta$ | $\lambda$ | $\zeta$ | $\lambda$ | $\zeta$ | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| JL | 500 | 0.021 | −4.22 | 0.051 | −6.54 | −0.025 | −8.70 | −0.032 | −11.2 | 0.959 | 1.5 |
|  | 1000 | ... | | ... | | ... | | 0.020 | −14.2 | 0.885 | 2.4 |
|  | 2000 | 0.050 | −5.97 | 0.098 | 3.14 | 0.051 | −11.8 | 0.020 | −18.1 | 0.951 | 1.8 |
| DO | 500 | 0.000 | −1.25 | 0.019 | −4.49 | −0.009 | −6.11 | −0.070 | −9.32 | 0.912 | 2.4 |
|  | 1000 | ... | | ... | | ... | | −0.081 | −9.96 | 0.657 | 4.1 |
|  | 2000 | 0.035 | −4.71 | 0.066 | −12.0 | 0.057 | −16.7 | 0.065 | −20.4 | 0.895 | 2.8 |
| BL | 500 | 0.012 | −2.15 | 0.052 | 3.35 | 0.013 | −7.66 | 0.055 | −9.53 | 0.952 | 1.8 |
|  | 1000 | ... | | ... | | ... | | 0.046 | −14.3 | 0.915 | 2.4 |
| FB | 2000 | 0.054 | −7.61 | 0.054 | −11.7 | 0.056 | −14.4 | 0.045 | −18.4 | 0.934 | 1.9 |
| DD | 500 | 0.024 | −0.904 | 0.055 | −2.61 | −0.012 | −3.08 | 0.126 | −6.61 | 0.955 | 1.4 |
| BZ | 1000 | ... | | ... | | ... | | −0.0005 | −15.1 | 0.888 | 2.4 |
| DD | 2000 | 0.050 | −3.07 | 0.041 | −6.18 | 0.045 | −12.3 | 0.037 | −11.3 | 0.943 | 2.1 |
| Mean | 500 | 0.031 | −2.06 | 0.044 | −4.22 | 0.004 | −6.08 | 0.034 | −9.76 | 0.978 | 1.3 |
|  | 1000 | ... | | ... | | ... | | 0.026 | −14.7 | 0.972 | 1.2 |
|  | 2000 | 0.055 | −6.59 | 0.065 | −8.58 | 0.055 | −13.6 | 0.045 | −17.2 | 0.966 | 1.6 |

the mean data, the proportion of variance accounted for by the two-parameter fit is 0.959, 0.972, and 0.897, respectively, for the 0.5-, 1.0-, and 2.0-kHz signals. Most of the residual variance can be attributed to measurement error. A second regression, in which all four parameters ($\beta_u$, $\beta_l$, $\alpha$, and $T_{max}$) were allowed to vary, indicated no significant or systematic departure from the values of the two-parameter fit. In particular, the slope values were equally often greater than and less than those of the two-parameter fit. The largest departures in mean slope values occurred for the 2.0-kHz signal: a $\beta_l$ of 47.9 (four-parameter fit) compared to 28.5 (two-parameter fit), and a $\beta_u$ of 155 compared to 143. In this worst case, the four-parameter fit accounted for an additional 4% of the total variance.

## B. Filter functions in simultaneous and forward masking as a function of masker level

Each of the panels of Fig. 2 gives simultaneous- and forward-masked thresholds for the 2.0-kHz signal obtained at a different masker level. The signal delay in forward masking was 5 ms. As before, the filter functions in simultaneous masking were estimated by selecting values of $\beta_u$, $\beta_l$, $\alpha$, and $T_{max}$ in Eq. (1) satisfying the least-squares criterion. As Table I shows, the four-parameter regression continues to account for a high proportion of the variance in the simultaneous-masked thresholds at the lower masker levels. The upward spread of masking with masker level is evident in the changing slopes of the filter functions. At the lowest masker
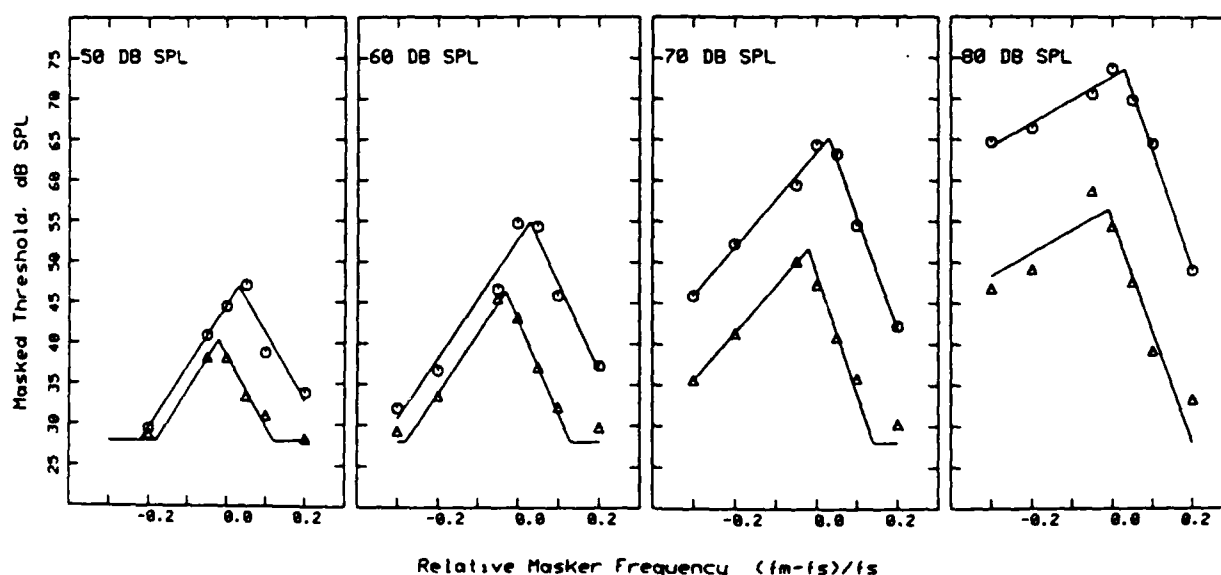


FIG. 2. Same as Fig. 1 except the simultaneous- and forward-masking filter functions are plotted for the 2000-Hz signal, at four masker levels: 50–80 dB SPL. Quiet threshold is 28 dB SPL and is designated by the knee in the filter functions at this level.

level, $\beta_u$ and $\beta_l$ are nearly equal; the filter function is roughly symmetric. As masker level grows, $\beta_u$ increases while $\beta_l$ decreases so that the filter function becomes highly asymmetric. Such changes in masking asymmetry are evident for all four subjects and replicate those commonly observed in simultaneous masking (e.g., Egan and Hake, 1950; Lutfi and Patterson, 1984; Patterson and Nimmo-Smith, 1980; Vogten, 1978a).

The forward-masking filter functions were derived as before from the simultaneous-masking filter functions using Eq. (2). The estimates of $\lambda$ and $\zeta$ for the individual and mean data are given in Table II. The two-parameter fits to the forward-masked thresholds are quite good. Excluding the 80-dB masker condition, which was described earlier, the forward-masking filter functions account for 97% or more of the total variability in the forward-masked thresholds at each level. This means that the level-dependent changes in masking asymmetry observed in simultaneous masking are maintained in forward masking. Note again that the estimates of $\lambda$ consistently place the MMF in forward masking slightly below that in simultaneous masking, and slightly below the frequency of the signal.

A similar pattern of results was obtained for the 0.5-kHz signal with the 5-ms signal delay. Figure 3 shows the data while Tables I and II give the parameters of the best-fitting filter functions. The forward-masking filter functions for the 0.5-kHz signal account for a comparably high proportion of the variability in both the individual and the mean forward-masked thresholds. For the mean data, the proportion of variance accounted for is 0.978. Changes in masking asymmetry with level are less pronounced for the 0.5-kHz signal than for the 2.0-kHz signal. At the lowest masker level, the filter function is asymmetric with $\beta_u$ less than $\beta_l$. This asymmetry is also opposite to that of the 2.0-kHz filter functions. Such reversals in masking asymmetry are not uncommon, particularly at low masker levels (e.g., Lutfi and Patterson, 1984; Zwicker and Jaroszewski, 1982). As before, $\beta_u$ increases with masker level while $\beta_l$ tends to decrease. Conse-

quently, the 0.5-kHz filter function becomes nearly symmetric at the highest masker level.

An important feature of both the 0.5- and 2.0-kHz data is the interaction that is observed between the effects of masker level and signal delay. At any given masker frequency, the threshold reduction that results from delaying the signal is greater at high masker levels than at low ones. For instance, when masker level is 50 dB, the difference between simultaneous- and forward-masked thresholds for on-frequency maskers (masker frequency equal to the signal frequency) is about 5 dB. When masker level is 80 dB, however, the dB difference is nearly quadrupled. The interaction between masker level and signal delay for on-frequency maskers has been described in detail by Jesteadt et al. (1982). The data of Figs. 2 and 3 indicate that the level–delay interaction behaves similarly for off-frequency maskers.

## C. Tuning curves in simultaneous and forward masking as a function of signal level

Figures 4 and 5 give tuning curves derived from the data of Figs. 2 and 3, respectively. The method for deriving these tuning curves follows that of Lutfi (1984) and Bacon and Viemeister (1985). Simultaneous- and forward-masking functions were estimated for each masker frequency by linear, least-squares regression of the mean thresholds on masker level. The masking functions were then used to compute the masker level at each frequency corresponding to a fixed threshold of 30, 40, 50, or 60 dB. Such point estimates based on the regression provide greater reliability than those based on a single mean provided that the relation between the variables is truly linear (Cohen and Cohen, 1975). Table III gives the obtained slope and intercept values. The estimated masking functions describe the data quite well. The worst case is represented by the function with the smallest slope (0.16), here the proportion of variance accounted for
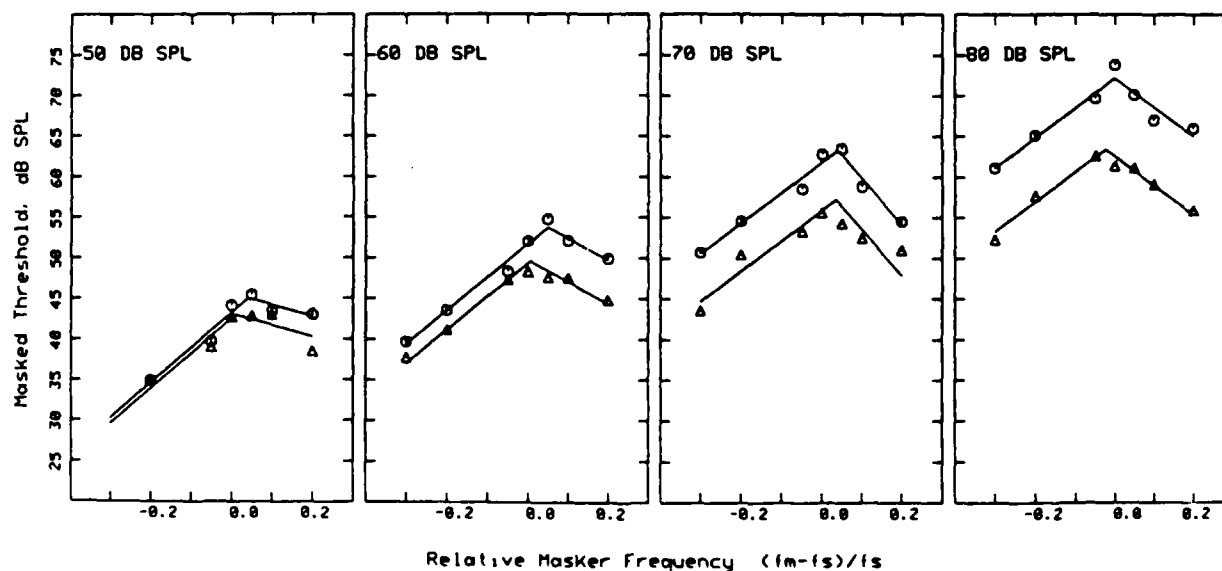


FIG. 3. Same as Fig. 2 except simultaneous- and forward-masking filter functions are shown for the 500-Hz signal.
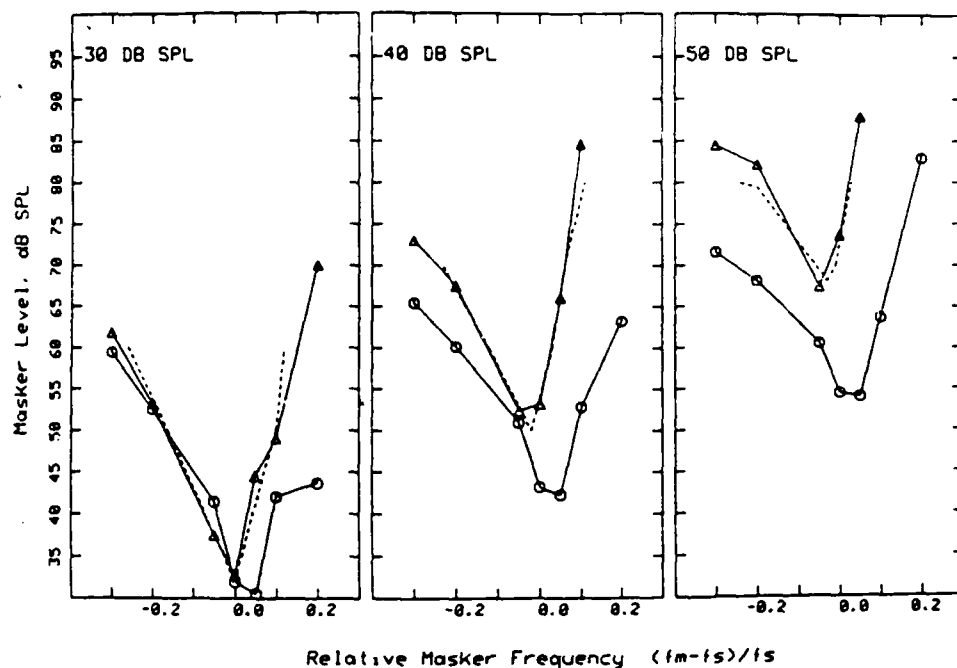
FIG. 4. Simultaneous- (circles) and forward (triangles) masking tuning curves at signal levels of 30–50 dB SPL. Signal frequency is 2000 Hz and signal delay is 5 ms. The obtained tuning curves (continuous lines) were derived from the masked thresholds of Fig. 2. The dashed lines are predictions derived from Eq. (2). See text for details.

was 91%. Also shown in Figs. 4 and 5 are predicted tuning curves (dotted lines). These curves were obtained by taking horizontal cuts through the predicted filter functions of Figs. 2 and 3. In some cases, it was also necessary to interpolate between points on the filter functions in order to determine the tip of the tuning curve.

The tuning curves obtained for the 2.0-kHz signal (Fig. 4) are representative of those reported in previous studies (e.g., Moore, 1978; Small, 1959; Vogten, 1978b; Weber, 1983). They have the familiar V shape in which the right branch of the V is quite steep and the left branch is slightly bowed. Note also that the tips of the tuning curves in forward masking are displaced slightly to the left of those is simultaneous masking. This disparity reflects the shifts in

the MMF noted earlier in the filter functions. Disparities between the tips of simultaneous- and forward-masking tuning curves have been reported previously by Vogten (1978a,b), Kidd and Feth (1981), Nelson and Freyman (1984), and Widin and Viemeister (1979a,b). In the study by Vogten, the MMF occurs at the signal frequency in forward masking, but slightly above the signal frequency in simultaneous masking. Vogten's tuning curves were obtained at low stimulus levels. The MMF shifts observed in the first and second panels of Fig. 4 replicate those reported at low levels. All of the remaining authors show that, as stimulus intensity is increased, the MMF in forward masking shifts to a frequency below that of the signal. This pattern of results is evident in the third panel of Fig. 4.
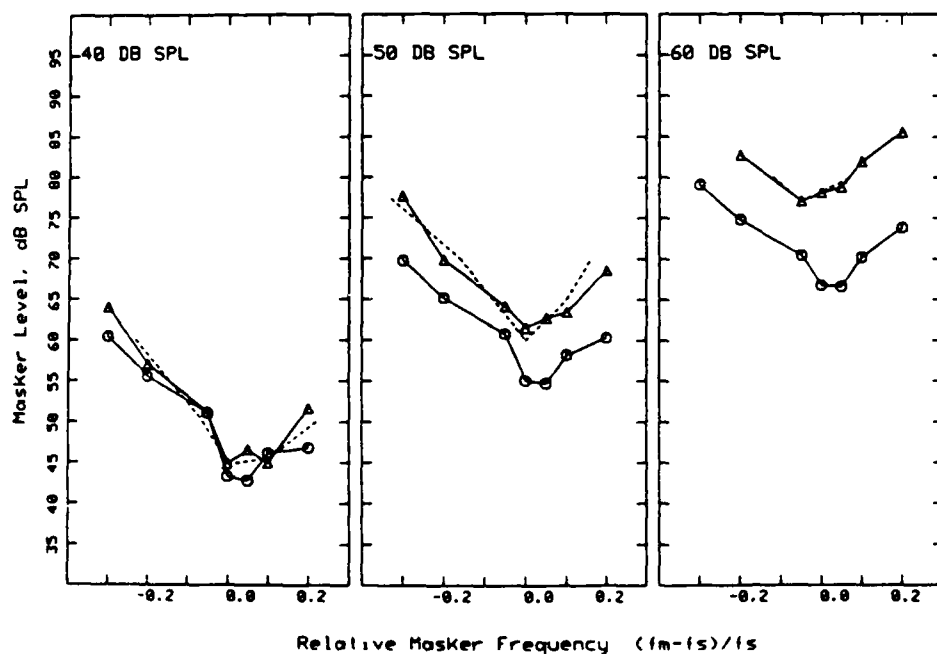


FIG. 5. Same as Fig. 4 except signal frequency is 500 Hz; the tuning curves were derived from the masked thresholds of Fig. 3.

| f Hz | t ms | g $(f - f_s)/f_s$ | Slope | Intercept dB SPL | n | $r^2$ |
|---|---|---|---|---|---|---|
| 500 | simultaneous | -0.30 | 1.07 | -24.6 | 3 | 1.000 |
| | | -0.20 | 1.04 | 17.8 | 5 | 0.998 |
| | | -0.05 | 1.03 | -12.6 | 5 | 0.998 |
| | | 0.00 | 0.85 | 3.2 | 6 | 0.984 |
| | | 0.05 | 0.84 | 4.0 | 5 | 0.997 |
| | | 0.10 | 0.83 | 1.7 | 5 | 0.995 |
| | | 0.20 | 0.74 | 5.4 | 4 | 0.967 |
| | 5 | -0.30 | 0.73 | -6.7 | 3 | 0.986 |
| | | -0.20 | 0.78 | -4.5 | 4 | 0.995 |
| | | -0.05 | 0.77 | 0.6 | 4 | 0.994 |
| | | 0.00 | 0.60 | 13.1 | 6 | 0.994 |
| | | 0.05 | 0.62 | 11.1 | 4 | 0.995 |
| | | 0.10 | 0.54 | 15.7 | 4 | 0.992 |
| | | 0.20 | 0.59 | 9.5 | 4 | 0.997 |
| 2000 | simultaneous | -0.30 | 1.63 | -66.7 | 3 | 0.992 |
| | | -0.20 | 1.28 | -37.2 | 5 | 0.989 |
| | | -0.05 | 1.05 | -13.5 | 5 | 0.990 |
| | | 0.00 | 0.89 | 1.5 | 6 | 0.996 |
| | | 0.05 | 0.85 | 4.1 | 5 | 0.992 |
| | | 0.10 | 0.92 | -8.6 | 5 | 0.994 |
| | | 0.20 | 0.51 | 7.7 | 4 | 0.976 |
| | 5 | -0.30 | 0.88 | -24.3 | 3 | 0.975 |
| | | -0.20 | 0.69 | -6.7 | 4 | 0.989 |
| | | -0.05 | 0.67 | 4.9 | 4 | 0.988 |
| | | 0.00 | 0.49 | 13.9 | 6 | 0.993 |
| | | 0.05 | 0.46 | 9.6 | 4 | 0.975 |
| | | 0.10 | 0.28 | 16.3 | 4 | 0.962 |
| | | 0.20 | 0.16 | 20.3 | 4 | 0.914 |
| | 10 | -0.20 | 0.80 | -14.3 | 4 | 0.998 |
| | | -0.05 | 0.47 | 10.2 | 4 | 1.000 |
| | | 0.00 | 0.42 | 17.0 | 6 | 0.996 |
| | | 0.05 | 0.36 | 11.7 | 4 | 0.999 |
| | | 0.10 | 0.40 | 7.9 | 4 | 0.953 |
| | 20 | -0.20 | 0.71 | -12.0 | 4 | 0.969 |
| | | -0.05 | 0.46 | 12.1 | 4 | 0.989 |
| | | 0.00 | 0.37 | 16.2 | 6 | 0.993 |
| | | 0.05 | 0.44 | 6.5 | 4 | 0.979 |
| | | 0.10 | 0.35 | 7.8 | 4 | 0.960 |
| | 40 | -0.05 | 0.43 | 8.6 | 4 | 0.986 |
| | | 0.00 | 0.19 | 22.8 | 6 | 0.927 |
| | | 0.05 | 0.31 | 11.5 | 4 | 0.980 |

The tuning curves of Fig. 4 are further typical in that overall they appear narrower in forward masking. For instance, for the 40-dB signal, the slope of the high-frequency branch of the tuning curve is roughly 190 dB/oct in simultaneous masking, while in forward masking it is near 320 dB/ oct. The respective slopes for the low-frequency branch of the tuning curve in simultaneous and forward masking are 40 and 45 dB/oct. These values are within the range of values that have been obtained in previous studies. The disparity of tuning in simultaneous and forward masking also appears to persist at high signal levels, consistent with the data of Moore and Glasberg (1986). Only the slope of the low-frequency tail of the forward-masking tuning curve (from $g = -0.3$ to $-0.2$) appears to become shallower at these high signal levels.

The disparity between the tuning curves in simultaneous and forward masking is related to the masker level–signal delay interaction described earlier. Recall that, at any given masker frequency, the threshold reduction that results from delaying the signal is greater at high masker levels than at low. This interaction affects the tuning curve because, for the tuning curve, masker level covaries with masker frequency. For the low-level, on-frequency maskers, the threshold reduction produced by delaying the signal is small. Thus the increment in masker level necessary to compensate for the threshold reduction is small. For the high-level, off-frequency maskers, the threshold reduction produced by delaying the signal is large; thus the corresponding increment in masker level is large. The fact that the tuning curves at 2.0-kHz are narrower in forward masking may, therefore, be understood in terms of a three-way interaction among the effects of masker frequency, masker level, and signal delay.

The tuning curves obtained for the 0.5-kHz signal are shown in Fig. 5. Unlike the tuning curves for the 2.0-kHz signal, these curves fail to evidence any significant difference in terms of the degree of apparent tuning in simultaneous and forward masking. Unfortunately, there are few comparable data in the literature at this low signal frequency. Vog-
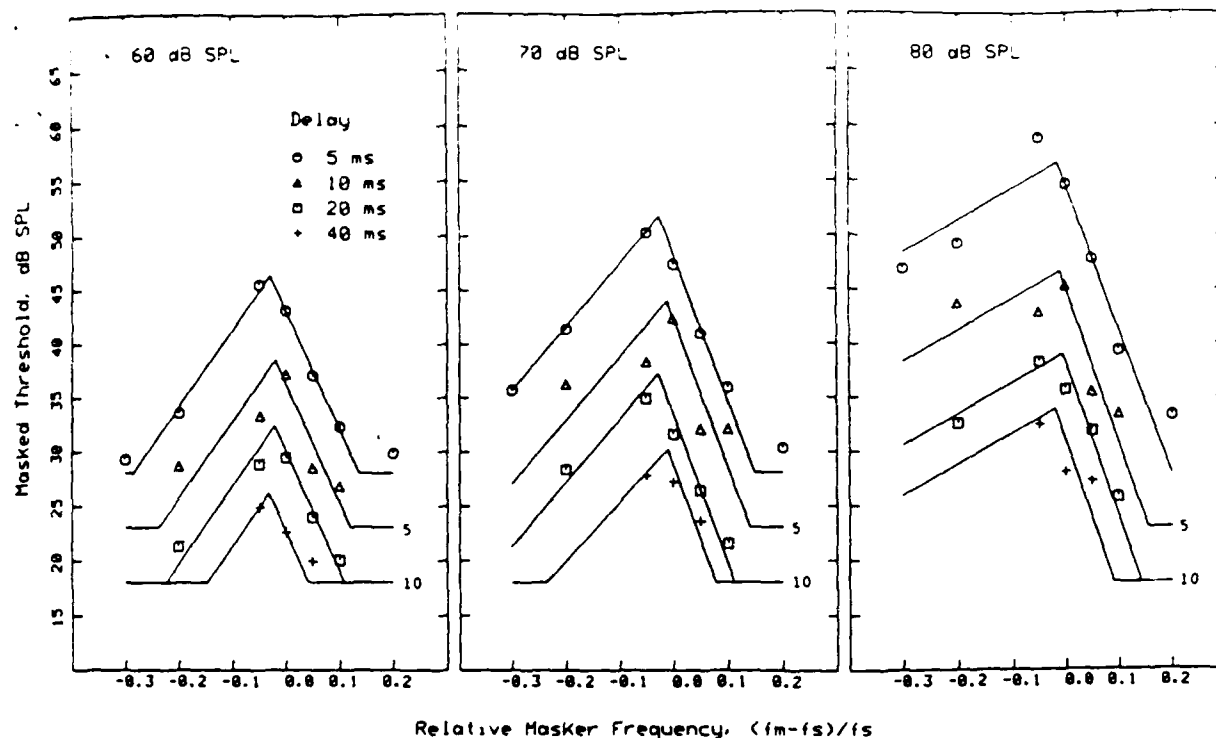
FIG. 6. Forward-masked thresholds for signal delays of 5 (circles). 10 (triangles), 20 (squares), and 40 (crosses) ms. Signal frequency is 2000 Hz. The forward-masking filter functions were derived from the simultaneous-masking filter functions according to Eq. (2). Note that the filter functions and corresponding thresholds for the 10-, 20-, and 40-ms delays have been shifted downward (number of dB indicated to the right of each function) to improve visible separation.

ten (1978b) reports tuning curves at 0.5 and 2.0 kHz for one subject. This subject's data agree with the present data inasmuch as the difference in tuning between simultaneous and forward masking was much less apparent for the 0.5-kHz signal. Moore *et al*. (1986) report comparable data from two subjects. One of these subjects showed narrower tuning curves in forward masking at 0.5 kHz, while the other failed to show any difference in tuning, at least within the range of frequencies used in the present study.

## D. Filter functions and tuning curves as a function of signal delay

To further test the generality of Eq. (2), we have obtained forward-masked thresholds for the 2.0-kHz signal as a function of signal delay. These forward-masked thresholds are given in Fig. 6. The filter functions drawn through these data were derived from the simultaneous-masking filter functions and Eq. (2) as before. Table IV gives the corre-

TABLE IV. Parameters $\lambda$ and $\zeta$ (entries) for forward-masking filter functions. Note that the $r$ value in the right-hand column refers to the proportion of total variability accounted for in all conditions of the corresponding row.

| Subject | Masker level dB SPL | $t = 5$ $\lambda$ | $\zeta$ | 10 $\lambda$ | $\zeta$ | 20 $\lambda$ | $\zeta$ | 40 ms $\lambda$ | $\zeta$ | rms error $r$ | dB |
|---------|------|-------|--------|-------|--------|-------|--------|-------|--------|-------|------|
| JL | 80 | 0.020 | − 18.1 | 0.037 | − 19.3 | 0.075 | − 26.7 | 0.040 | − 29.2 | 0.931 | 1.9 |
|    | 70 | 0.051 | − 11.8 | 0.077 | − 11.8 | 0.067 | − 17.1 | 0.059 | − 22.4 | 0.915 | 2.1 |
|    | 60 | 0.098 | 3.1 | 0.082 | − 9.1 | 0.060 | − 12.6 | 0.060 | − 20.2 | 0.927 | 1.5 |
| DO | 80 | 0.065 | − 20.4 | 0.065 | − 25.5 | 0.045 | − 28.3 | 0.074 | − 35.2 | 0.671 | 3.8 |
|    | 70 | 0.057 | − 16.7 | 0.062 | − 21.8 | 0.021 | − 24.8 | 0.025 | − 30.1 | 0.641 | 3.0 |
|    | 60 | 0.066 | − 12.0 | 0.077 | − 16.2 | 0.086 | − 16.4 | 0.106 | − 20.2 | 0.840 | 2.0 |
| FB | 80 | 0.045 | − 18.4 | 0.032 | − 24.7 | 0.039 | − 23.4 | 0.050 | − 29.5 | 0.805 | 3.2 |
|    | 70 | 0.056 | − 14.4 | 0.025 | − 17.0 | 0.044 | − 20.4 | 0.090 | − 19.7 | 0.914 | 1.7 |
|    | 60 | 0.054 | − 11.7 | 0.022 | − 16.6 | 0.027 | − 16.0 | 0.072 | − 20.0 | 0.823 | 2.1 |
| DD | 80 | 0.037 | − 11.3 | 0.045 | − 15.8 | 0.037 | − 21.1 | 0.006 | − 28.2 | 0.864 | 2.4 |
|    | 70 | 0.045 | − 12.3 | 0.037 | − 16.4 | 0.052 | − 16.3 | 0.020 | − 23.9 | 0.851 | 1.8 |
|    | 60 | 0.041 | − 6.2 | 0.035 | − 8.0 | 0.036 | − 10.0 | 0.040 | − 14.7 | 0.864 | 1.6 |
| Mean | 80 | 0.045 | − 17.2 | 0.040 | − 22.1 | 0.035 | − 24.7 | 0.050 | − 29.8 | 0.871 | 2.2 |
|    | 70 | 0.055 | − 13.6 | 0.041 | − 16.4 | 0.057 | − 18.1 | 0.040 | − 25.1 | 0.879 | 1.7 |
|    | 60 | 0.065 | − 8.6 | 0.055 | − 11.4 | 0.057 | − 12.5 | 0.067 | − 18.8 | 0.901 | 1.6 |

sponding values of $\lambda$ and $\zeta$, and the proportion of variance accounted for. Note that the threshold and filter functions are shifted downward with each delay to improve visible separation. The data described previously for the 5-ms delay are also included for comparison. Individually, the proportion of variance accounted for by the two-parameter fits tends to be lower at longer signal delays. For subject DO, in particular, the predicted functions account for less than 70% of the total variability in some cases. There were two factors that contributed to the lower proportion of variance accounted for individually. First, the range over which masking could be measured at the long delays was restricted. Consequently, there were fewer thresholds and the total variability among thresholds was smaller. Second, at the long signal delays, the individual off-frequency thresholds tended to be more variable so that there was greater measurement error. This was most true of subject DO.

Turning to the mean data, the two-parameter fits account, respectively, for 90%, 88%, and 87% of the variability in the mean forward-masked thresholds at masker levels of 60–80 dB. In each case, the rms error between the predicted and obtained thresholds is close to 2 dB. The corresponding four-parameter fits accounted, respectively, for 96%, 95%, and 94% of the variability. The slopes of the filter functions resulting from the four-parameter fits at the longer signal delays cannot be considered very reliable given the limited number of data points defining each curve. However,

in one case, these slopes did consistently and significantly deviate from those predicted by the two-parameter fit. For the 10-ms signal delay, the slopes of the four-parameter fits on the low-frequency side were generally smaller, ranging from 36.0 dB for the 60-dB level masker to 0.2 dB for the 80-dB level masker. The corresponding slopes for the two-parameter fits (see Table I) range from 72.0–28.5 dB.

The first panel of Fig. 7 gives examples of tuning curves corresponding to these data. Predictions are shown as dotted lines as before. The 35-dB signal in this case was selected to be representative of the low-level signals for which tuning curves are most frequently reported in the literature. The major effect of increasing signal delay is an overall elevation of the tuning curve accompanied by a shift in the the tip of the curve (the point at the MMF) to a frequency below that of the signal. Kidd and Feth (1981) and Nelson and Freyman (1984) report identical shifts in the tips of tuning curves with increasing delay. In their data, the shift in tips is often accompanied by a decrease in the slope of the low-frequency branch of the tuning curve, as is evident in the predicted curves of Fig. 7. Nelson and Freyman (1984) report that the differences among their tuning curves are largely eliminated if signal level–signal delay combinations are selected such that the level of maskers near the tips of the curves are equated. When this is done, their tuning curves nearly superimpose. The second panel of Fig. 7 shows tuning curves in which the level of maskers near the tips of the
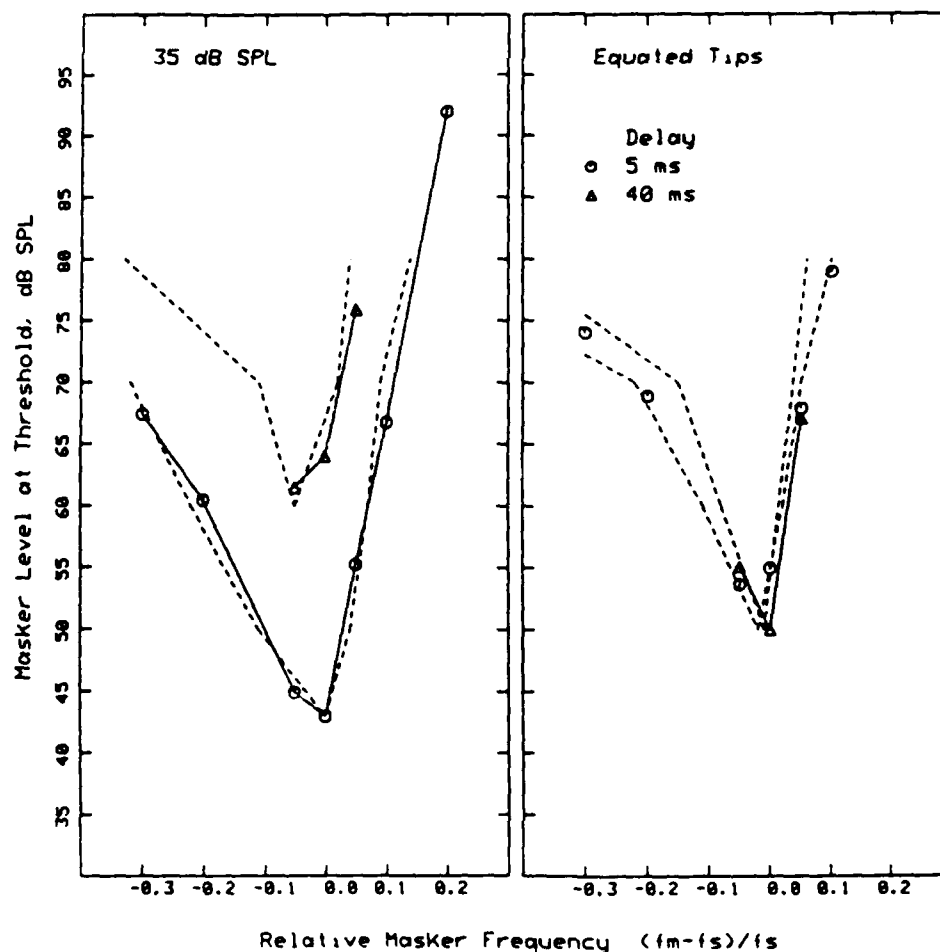


FIG. 7. Left panel: forward-masking tuning curves for a 35-dB SPL signal at signal delays of 5 (circles) and 40 (triangles) ms. The tuning curves were derived from the data of Fig. 6 (see text for details). Dashed lines represent predictions based on Eq. (2). Right panel: Signal level is selected for each signal delay so as to equate masker level near the tips of the tuning curves. The lower of the two dashed curves gives the predictions for the 5-ms delay; the solid curve is omitted for clarity of presentation.

curves has been equated. For the signal delays of 5 and 40 ms. the corresponding signal levels are 40.8 and 32.3 dB SPL. In keeping with the data of Nelson and Freyman (1984), the differences among these curves are largely reduced

## E. Analyses with fewer parameters

Earlier data of Lutfi (1984) were described well by assuming a $\lambda$ of 0. The value of $\zeta$ was then given as the dB difference between thresholds in on-frequency $(g = 0)$ forward and on-frequency simultaneous masking. Specifically, $\zeta = G - H(0)$, where, $G$ is the on-frequency, forward-masked threshold. Substituting into Eq. (2), the earlier model's predictions were given by the relation

$$F(g) = H(g) + [G - H(0)]. \tag{3}$$

Note that this model contains no free parameters. In terms of Eq. (2), $\lambda$ is a predetermined constant and $\zeta$ is extracted directly from the data. The fixed-parameter model predicts that for a fixed-level masker, the time decay of masking in dB is the same for all masker frequencies and is given by the difference $G - H(0)$.

In this section, the earlier model and two modifications are applied to the present data. In the first version, the time decay of masking is assumed to be constant as before, but no specific relation is assumed between masker level and the

amount of decay. This model's predictions are given by

$$F(g) = H(g) - \zeta. \tag{4}$$

where $\zeta$ is allowed to vary as a free parameter. In the second version, the time decay of masking is assumed to be constant with a shift in frequency equal to the shift in the MMF. This model is similar to the two parameter model with the exception that the shift in the MMF was estimated by allowing $\lambda$ to vary with the constraint $\zeta = G - H(\lambda)$. The variable-$\lambda$ model's predictions are given by the relation

$$F(g) = H(g + \lambda) + [G - H(\lambda)]. \tag{5}$$

The results of the regressions for all three models are shown in Table V. The proportion of variance accounted for by four-parameter fits to the data is included for comparison. The fixed-parameter model clearly does a poor job of summarizing the data. On average, the proportion of variance accounted for is over 18% less than when all four parameters are allowed to vary. The variable-$\zeta$ model does only slightly better. On average, the proportion of variance it accounts for is 14% less than when all four parameters are allowed to vary. Only the variable-$\lambda$ model provides a comparatively good fit to the data. It misses on average less than 3% of the total variability accounted for by the four-parameter model. Figure 8 summarizes, for all conditions of this study, the forward-masked thresholds and the corresponding predictions of the variable-$\lambda$ model. The dashed lines correspond to points of equality between the obtained and

TABLE V. Parameters for forward masking filter functions resulting from three assumed models (see text for a full explanation). The regression in each case was performed on the mean data. Bottom rows give the proportion variance accounted for by each model and the corresponding rms error at each signal frequency. The regression results for the four-parameter fits are included for comparison. The difference between the $\alpha$ values obtained in simultaneous and forward masking for the four-parameter fit provides an independent estimate of the frequency shift $\lambda$.

| $f$, Hz | $t$ ms | Masker level dB SPL | Fixed-parameter $\lambda$ | Fixed-parameter $\zeta$ | Variable-$\zeta$ $\lambda$ | Variable-$\zeta$ $\zeta$ | Variable-$\lambda$ $\lambda$ | Variable-$\lambda$ $\zeta$ | Four-parameter $\lambda = \alpha_s - \alpha_f$ |
|---|---|---|---|---|---|---|---|---|---|
| 500 | 5 | 80 | 0.000 | −10.1 | 0.000 | −8.93 | 0.045 | −10.1 | 0.045 |
| | | 70 | 0.000 | −7.60 | 0.000 | −6.10 | −0.005 | −7.60 | 0.040 |
| | | 60 | 0.000 | −5.10 | 0.000 | −3.68 | 0.037 | −5.10 | 0.059 |
| | | 50 | 0.000 | −2.60 | 0.000 | −1.61 | 0.042 | −2.60 | −0.004 |
| 1000 | 5 | 80 | 0.000 | −17.4 | 0.000 | −15.7 | 0.022 | −17.4 | 0.019 |
| 2000 | 5 | 80 | 0.000 | −19.6 | 0.000 | −18.6 | 0.052 | −19.6 | 0.056 |
| | | 70 | 0.000 | −15.6 | 0.000 | −14.8 | 0.055 | −15.6 | 0.062 |
| | | 60 | 0.000 | −11.6 | 0.000 | −9.75 | 0.065 | −11.6 | 0.069 |
| | | 50 | 0.000 | −7.60 | 0.000 | −7.73 | 0.055 | −7.60 | 0.062 |
| | 10 | 80 | 0.000 | −22.1 | 0.000 | −24.0 | 0.037 | −22.1 | 0.030 |
| | | 70 | 0.000 | −17.4 | 0.000 | −17.6 | 0.045 | −17.4 | 0.030 |
| | | 60 | 0.000 | −12.7 | 0.000 | −14.1 | 0.050 | −12.7 | 0.035 |
| | 20 | 80 | 0.000 | −26.9 | 0.000 | −26.3 | 0.007 | −26.9 | 0.057 |
| | | 70 | 0.000 | −21.7 | 0.000 | −20.3 | 0.065 | −21.7 | 0.066 |
| | | 60 | 0.000 | −16.5 | 0.000 | −15.0 | 0.067 | −16.5 | 0.049 |
| | 40 | 80 | 0.000 | −34.7 | 0.000 | −32.4 | −0.007 | −34.7 | ⋯ |
| | | 70 | 0.000 | −27.7 | 0.000 | −26.4 | 0.025 | −27.7 | ⋯ |
| | | 60 | 0.000 | −20.7 | 0.000 | −19.1 | 0.060 | −20.7 | ⋯ |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 500 | | Proportion of | 0.924 | | 0.962 | | 0.969 | | 0.992 |
| 1000 | | variance | 0.861 | | 0.888 | | 0.968 | | 0.977 |
| 2000 | | accounted for | 0.616 | | 0.661 | | 0.909 | | 0.971 |

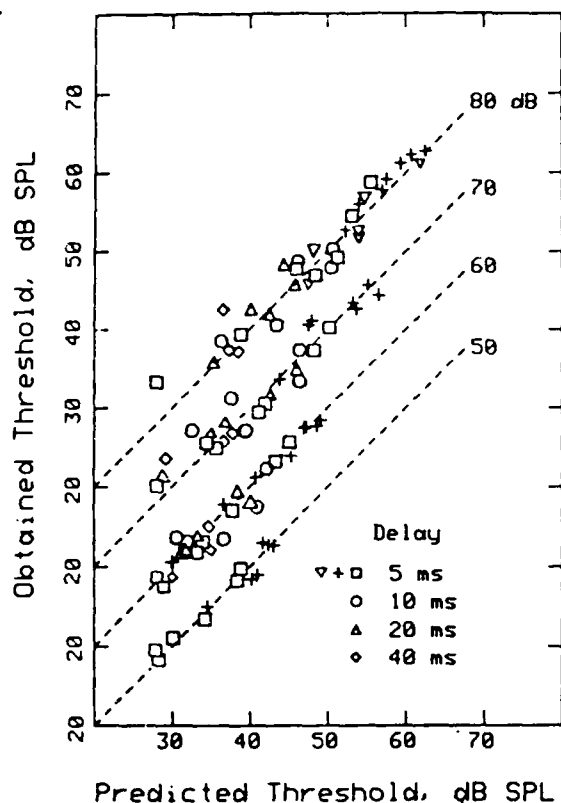| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 500 | | rms | 2.1 | | 1.5 | | 1.3 | | 0.7 |
| 1000 | | error | 1.0 | | 0.9 | | 0.4 | | 0.4 |
| 2000 | | in dB | 4.2 | | 3.9 | | 2.0 | | 1.1 |

FIG. 8. Predictions of the variable-$\lambda$ model. Forward-masked thresholds are shown for all conditions of this study. The dashed lines correspond to points of equality between the obtained and the predicted thresholds. They have been displaced from one another by 10 dB to allow comparisons at each masker level (parameter). The different symbols denote the different signal delay masker frequencies appear as repeated symbols for each masker level and signal delays; combination. For the 5-ms delay, the signal frequencies were 500 (crosses), 1000 (inverted triangles), and 2000 (squares) Hz. In all other cases, the signal frequency is 2000 Hz.

the predicted thresholds. They have been displaced from one another by 10 dB to allow comparisons at each masker level. The different symbols denote signal delay; masker frequencies and signal frequencies appear as repeated symbols for each masker level–signal delay combination. The correspondence between the obtained and predicted thresholds is

good. Overall the one-parameter, variable-$\lambda$ model accounts, respectively, for 97%, 97%, and 91% of the variance in the forward-masked thresholds for the 0.5-, 1.0-, and 2.0-kHz signals. The corresponding rms error is 1.3, 0.4, and 2.0 dB.

## III. DISCUSSION

### A. The interaction between masker level and signal delay

It has long been known that the time decay of masking is related to the overall level of the masker (Gardner, 1947; Harris et al., 1951; Plomp, 1964; Zwislocki et al., 1959). Perhaps, the most comprehensive data on this relation are those of Jesteadt et al. (1982). These authors have shown that the family of functions relating signal threshold to signal delay at each masker level converge at a common delay. The rate of masking decay is greater at high masker levels than at low. Widin and Viemeister (1979r) interpret the level–delay interaction to reflect the dependence of masking decay on the overall level of the masker. However, an alternative interpretation is that the interaction reflects the dependence of masking decay on the initial amount of masking (Moore and Glasberg, 1981). In the earlier studies, as well as the study by Jesteadt et al., the overall level of the masker is correlated with the initial amount of masking at some short delay.

The data of the present study bear on this issue. For the filter function, the initial amount of masking (simultaneous masking) decreases with increasing frequency separation between signal and masker, but the overall level of the masker remains constant. Thus, if masking decay depends on the initial amount of masking, simultaneous- and forward-masking filter functions should be nonparallel; they should be parallel if masking decay depends on the overall level of the masker. The present data tend to support the conclusion of Widin and Viemeister that masking decay depends on the overall level of the masker. Based on their own data, however, Moore and Glasberg (1981) came to the opposite conclusion. They show filter functions (their Fig. 8), obtained
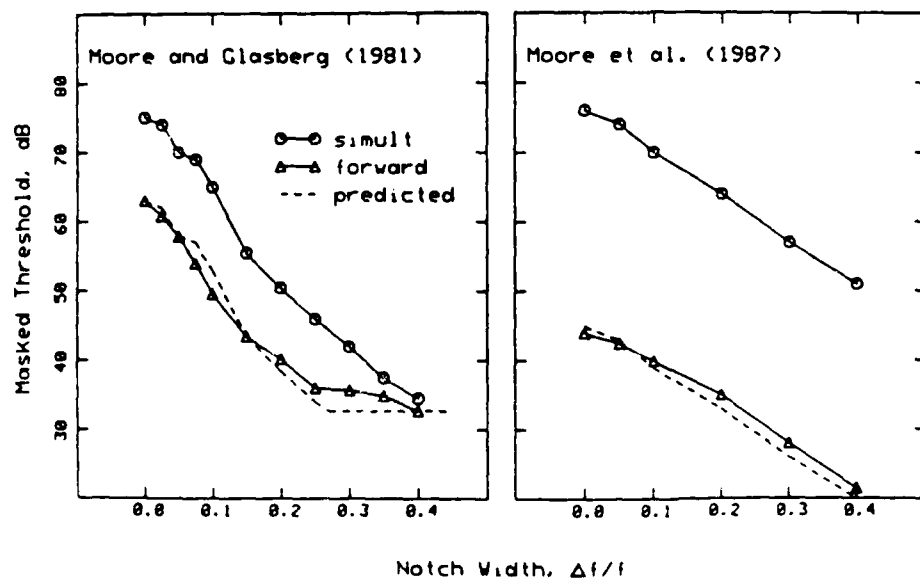


FIG. 9. Left panel: simultaneous- (circles) and forward- (triangles) masked thresholds from the study of Moore and Glasberg (1981), subject IB. Signal duration is 5 ms. Right panel: simultaneous- and foward-masked thresholds from the study of Moore et al. (1987). In each panel, the dashed line gives the prediction of the variable-$\lambda$ model for $\lambda$ equal to 0.

using a notched-noise masker, which are slightly sharper in forward masking.

Part of the reason for this apparent discrepancy has to do with the way in which Moore and Glasberg present their data. The forward-masked thresholds shown in their Fig. 8 are not the actual masked thresholds, but rather a transformation of these thresholds based on a broadband noise masking condition. We have performed the inverse transformation to allow comparison between the simultaneous- and forward-masked thresholds actually obtained. The data shown in Fig. 9 are from a representative subject (IB). The dashed line gives the prediction of Eq. (5) with $\lambda$ equal to 0; in this case, $H$ is equivalent to the curve drawn through the simultaneous-masked thresholds. Quiet threshold were not reported for these subjects; therefore, the horizontal portion of predicted curve corresponds to an estimate of quiet threshold. The obtained filter function in forward masking does appear to be shallower than that predicted by Eq. (5). The discrepancy is so small, however, that it is not possible to decide based on these data alone whether the decay of masking ultimately depends on overall masker level or on the initial amount of masking. A stronger test requires measuring the forward-masking filter functions at longer signal delays. Moore and Glasberg (1981) report a progressive broadening of forward-masking filter functions with signal duration that is related to signal delay. Unfortunately, Eq. (5) cannot be applied to these data as the corresponding simultaneous-masking filter functions are missing. Equation (5) can, however, be applied to recent data of Moore et al. (1987) where simultaneous- and forward-masking filter functions are reported for a longer duration signal. The data from a representative subject (FL) and the prediction of Eq. (5) are shown in the second panel of Fig. 9. The filter functions, which were obtained using a 20-ms signal, are shallower than those shown in the first panel of this figure where the signal duration is 5 ms. This result is consistent with the data of Moore and Glasberg. Once again, the deviation from the predicted curve is quite small.

Although the present data might appear to support the conclusion that masker level, not initial amount of masking, is the critical variable in determing the rate of masking decay, it is important to note that the two interpretations need not be mutually exclusive. Small deviations from parallel filter functions may indicate a weak relation to initial amount of masking, although a larger proportion of the variance may be accounted for by the relation to masker level. However one chooses to interpret the interaction, it is clear from the present as well as the past studies that the decay of masking is generally greater at higher masker levels.

## B. Implications for interpreting measures of frequency selectivity

It is reasonable to suspect that certain procedures may yield differences between simultaneous and forward masking that have little to do with any "real" difference in auditory frequency selectivity. This concern was intimated early on (Widin and Viemeister, 1979a; Wightman et al., 1977). The present study underscores this concern inasmuch as the disparity between simultaneous- and forward-masking tun-

ing curves is attributed to an interaction between masker level and signal delay that is peculiar to the tuning curve procedure. It is difficult to determine to what extent the level–delay interaction might have affected differences among tuning curves observed in past studies. These studies have not, in general, obtained all the measures within a study necessary to evaluate the level–delay interaction independently of observed frequency effects. Special care is required, therefore, in interpreting the previously observed differences between simultaneous and forward masking. This conclusion pertains as well to the broadening of tuning curves observed at longer signal delays (Kidd and Feth, 1981; Nelson and Freyman, 1984). It is known, for instance, that filter functions begin to broaden at high masker levels (Patterson, 1971; Lutfi and Patterson, 1984; Weber, 1977). Since high-level maskers are typically required to mask signals at long delays, one may expect that tuning curves would also broaden at long signal delays. As Nelson and Freyman (1984) showed, such effects can be compensated for by equating masker levels near the tips of the tuning curves. Of course such problems can be avoided by fixing masker level, and varying signal level to threshold, as in the filter function procedure. This procedure has the advantage of allowing an estimate of shape of the auditory filter at each masker level (e.g., Lutfi and Patterson, 1984).

A related measure in which masker level is fixed is the masking pattern. The masking pattern differs from the filter function in that signal frequency rather than masker frequency is plotted along the abscissa; masker frequency is fixed. For the masking pattern, the level–delay interaction must be assessed by measuring the on-frequency forward-masking function $G$ at each signal frequency. Although the effects are small, the on-frequency forward-masking function can vary with signal frequency (e.g., Jesteadt et al., 1982). Thus, according to Eq. (5), masking patterns need not be parallel with signal delay, unlike the filter functions (cf. Lutfi, 1985; Moore, 1985). Kidd and Feth (1982) report masking patterns as a function of signal delay for a 1.0-kHz, sinusoidal masker, and, indeed, the high-frequency branches of their masking patterns become more shallow with increasing delay. Unfortunately, Kidd and Feth do not present the corresponding on-frequency masking functions that would be required to apply the predictions of Eq. (5) to their data.

## C. Additivity and the MMF

To predict the forward-masked thresholds of the present study, we required the estimation of the parameter $\lambda$ equal to the shift in the MMF from simultaneous to forward masking. Consider, however, the special case in which $\lambda$ equals 0 (the fixed-parameter model). Then, Eq. (5) may be rewritten in the form,

$$F(g) = G + W(g), \tag{6}$$

where $W(g) \simeq H(g) - H(0)$ is the dB attenuation characteristic of the auditory filter (see Patterson, 1976). Equation (6) suggests a fundamental property of masking. That is, if we identify $G$ with the time decay of masking and $W$ with the frequency selectivity of the system, then the implication of

Eq. (6) is that these two processes are independent and additive (additive, that is, in dB). Of course, as previously shown, the present data are poorly summarized when $\lambda$ is assumed to be 0. Therefore, the relation implied by Eq. (6) is either false or some third unrelated factor is responsible for the shift in the MMF in forward masking.

McFadden (1986) has recently reviewed arguments favoring the latter possibility. The most compelling interpretation attributes the shift in the MMF to a basalward displacement of the traveling-wave envelope with increasing stimulus intensity. According to McFadden, when stimulus intensity is high, the region maximally "fatigued" along the cochlear partition occurs basal to the region maximally fatigued when stimulus intensity is low. Thus, at high intensities, signals at frequencies above the masker would be expected to undergo the greatest amount of masking. McFadden cites physiological evidence for a basalward displacement with intensity which lends credence to this interpretation (e.g., Evans, 1977; Rhode, 1978; Russell and Sellick, 1978; Sachs and Abbas, 1974). He further argues that the greatest effects should be observed in forward masking, as in the present study, where certain confounding interactions between signal and masker are eliminated. Other interpretations have been proposed by Vogten (1978b) and Widin and Viemeister (1979b). Vogten describes the shift in the MMF observed at low stimulus intensities in terms of suppression, while Widin and Viemeister emphasize the importance in forward masking of the short-term spectrum at the offset of the masker. Of all of these interpretations, only Vogten's is clearly inconsistent with the additive relation implied by Eq. (6).

## D. Individual differences in tuning curves

The decision to average the data of the four subjects in each condition of this study was based on the similarity among the individual filter functions in each condition. In other words, the thresholds overall did not differ significantly across subjects. When the data were replotted as tuning curves (constant threshold), however, differences among subjects were greatly exaggerated, particularly on the high-frequency side of the tuning curve, and particularly in forward masking. The reason for this is clear. Consider the slope values of the masking functions in Table III. The slope values are generally quite small for forward maskers with frequencies above that of the signal ($g > 0$). For instance, when $g = 0.2$, the slope of the masking function for the 2.0-kHz signal presented at the 5-ms delay is only 0.16. This means that a 1.6-dB difference in threshold in this condition corresponds to a 10-dB difference in masker level for the tuning curve. Quite frequently, such small differences in threshold can be expected to produce rather large differences in psychophysical tuning curves. Consequently, special care may be required in interpreting individual differences among tuning curves. This is particularly true in studies where tuning curves are reported for only one or two subjects, or where comparisons are made between the tuning curves of individual normal-hearing and hearing-impaired subjects.

## IV. SUMMARY

The major results of this study can be summarized as follows: (1) Simultaneous- and forward-masked thresholds are described well by parallel filter functions; (2) maximum forward masking typically occurs at a frequency below that of maximum simultaneous masking; (3) except, perhaps, at low signal frequencies, forward-masking tuning curves are narrower than simultaneous-masking tuning curves, even at high signal levels; (4) differences among forward-masking tuning curves are largely eliminated when signal-level, signal-delay combinations are chosen to equate masker levels near the tips of the tuning curves; (5) a single frequency-selective function estimated exclusively from the simultaneous-masked thresholds can be used to predict results (1)–(4). The results imply that a single frequency-selective process can account for commonly observed differences between simultaneous- and forward-masked measures of frequency selectivity.

[1]The function $H$ is actually proportional to the integral of the auditory filter characteristic evaluated over the frequency band of the masker. However, because the frequency band of the masker is relatively small, direct comparisons between the bandwidths of $H$ and the filter characteristic are permissible.

Bacon, S. P., and Moore, B. C. J. (1986). "Temporal effects in masking and their influence on psychophysical tuning curves," J. Acoust. Soc. Am. **80**, 1638–1645.

Bacon, S. P., and Viemeister, N. F. (1985). "Simultaneous masking by gated and continuous sinusoidal maskers," J. Acoust. Soc. Am. **78**, 1220–1230.

Barnet, V., and Lewis, T. (1978). *Outliers in Statistical Data* (Wiley, New York).

Childers, D., and Durling, A. (1975). *Digital Filtering and Signal Processing* (West, St. Paul, MN), pp. 437–439.

Cohen, J., and Cohen, P. (1975). *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences* (Erlbaum, Hillsdale, NJ).

Duifhuis, H. (1976). "Cochlear nonlinearity and second filter: Possible mechanism and implications," J. Acoust. Soc. Am. **59**, 408–423.

Egan, J. P., and Hake, H. W. (1950). "On the masking pattern of a simple auditory stimulus," J. Acoust. Soc. Am. **22**, 622–630.

Ehmer, R. H., and Ehmer, B. J. (1969). "Frequency pattern of residual masking by pure tones measured on the Békésy audiometer," J. Acoust. Soc. Am. **46**, 1445–1448.

Evans, E. F. (1975). "Auditory frequency selectivity and the cochlear nerve" in *Facts and Models in Hearing*, edited by E. Zwicker and E. Terhardt (Springer, New York), pp. 118–129.

Evans, E. F. (1977). "Frequency Selectivity at high signal levels of single units in cochlear nerve and nucleus," in *Psychophysics and Physiology of Hearing*, edited by E. F. Evans and J. P. Wilson (Academic, New York), pp. 185–192.

Gardner, M. B. (1947). "Short duration auditory fatigue as a method of

classifying hearing impairment." J. Acoust. Soc. Am. 19, 178-190

Greenwood, D. D. (1971). "Aural combination tones and auditory masking." J. Acoust. Soc. Am. 50, 502-543.

Harris, J. D., Rawnsley, A. I., and Kelsey, P. (1951). "Studies in short-duration auditory fatigue: I. Frequency differences as a function of intensity." J. Exp. Psychol. 42, 430-436.

Houtgast, T. (1972). "Psychophysical evidence for lateral inhibition in hearing." J. Acoust. Soc. Am. 51, 1885-1894.

Houtgast, T. (1974). "Lateral suppression in hearing: A psychophysical study of the ear's capability to preserve and enhance spectral contrasts." Doctoral dissertation, Academishe Pers. B. V., Amsterdam.

Jesteadt, W., Bacon, S. P., and Lehman, J. R. (1982). "Forward masking as a function of frequency, masker level, and signal delay." J. Acoust. Soc. Am. 71, 950-962.

Jesteadt, W., and Norton, S. J. (1985). "The role of suppression in psychophysical measures of frequency selectivity." J. Acoust. Soc. Am. 78, 365-374.

Kiang, N. Y.-S., Watanabe, T., Thomas, E. C., and Clark, L. F. (1965). "Discharge patterns of single fibers in the cat's auditory nerve." Res. Monogr. 35 (MIT, Cambridge, MA).

Kidd, G., Jr., and Feth, L. L. (1981). "Patterns of residual masking." Hear. Res. 5, 49-67.

Levitt, H. (1971). "Transformed up-down methods in psychoacoustics." J. Acoust. Soc. Am. 49, 467-477.

Lutfi, R. A. (1984). "Predicting frequency selectivity in forward masking from simultaneous masking." J. Acoust. Soc. Am. 76, 1045-1050.

Lutfi, R. A. (1985). "A reply to comments by Moore [J. Acoust. Soc. Am. 78, 253-255 (1985)]." J. Acoust. Soc. Am. 78, 255-257.

Lutfi, R. A., and Patterson, R. D. (1984). "On the growth of masking asymmetry with stimulus intensity." J. Acoust. Soc. Am. 76, 739-745.

McFadden, D. (1986). "The curious half-octave shift: Evidence for a basalward migration of the traveling-wave envelope with increasing intensity." in Basic and Applied Aspects of Noise-Induced Hearing Loss, edited by R. J. Salvi, D. Henderson, R. P. Hamerick, and V. Colletti (Plenum, New York).

Moore, B. C. J. (1978). "Psychophysical tuning curves measured in simultaneous and forward masking." J. Acoust. Soc. Am. 63, 524-532.

Moore, B. C. J. (1985). "Comments on 'Predicting frequency selectivity in forward masking from simultaneous masking' [J. Acoust. Soc. Am. 76, 1045-1050 (1984)]." J. Acoust. Soc. Am. 78, 253-255.

Moore, B. C. J., and Glasberg, B. R. (1981). "Auditory filter shapes derived in simultaneous and forward masking." J. Acoust. Soc. Am. 70, 1003-1014.

Moore, B. C. J., and Glasberg, B. R. (1986). "Comparisons of frequency selectivity in simultaneous and forward masking for subjects with unilateral cochlear impairments." J. Acoust. Soc. Am. 80, 93-107.

Moore, B. C. J., Glasberg, B. R., and Roberts, B. (1984). "Refining the measurement of psychophysical tuning curves." J. Acoust. Soc. Am. 76, 1057-1066.

Moore, B. C. J., Poon, W. F., Bacon, S. P., and Glasberg, B. R. (1987). "The temporal course of masking and the auditory filter shape." J. Acoust. Soc. Am. 81, 1873-1880.

Munson, W. A., and Gardner, M. B. (1950). "Loudness patterns: A new approach." J. Acoust. Soc. Am. 22, 177-190.

Nelson, D. A., and Freyman, R. L. (1984). "Broadened forward-masked

tuning curves from intense masking tones: Delay-time and probe level manipulations." J. Acoust. Soc. Am. 75, 1570-1577.

O'Loughlin, B. J., and Moore, B. C. J. (1981). "Off-frequency listening: Effects on psychophysical tuning curves obtained in simultaneous and forward masking." J. Acoust. Soc. Am. 69, 1119-1125.

Patterson, R. D. (1971). "Effect of amplitude on the auditory filter shape." J. Acoust. Soc. Am. 49, 81.

Patterson, R. D. (1976). "Auditory filter shapes derived with noise stimuli." J. Acoust. Soc. Am. 59, 640-654.

Patterson, R. D., and Nimmo-Smith, I. (1980). "Off-frequency listening and auditory-filter asymmetry." J. Acoust. Soc. Am. 67, 229-245.

Plomp, R. (1964). "Rate of decay of auditory sensation." J. Acoust. Soc. Am. 36, 277-282.

Rhode, W. S. (1978). "Some observations on cochlear mechanics." J. Acoust. Soc. Am. 64, 158-176.

Russell, I. J., and Sellick, P. M. (1978). "Intracellular studies of hair cells in the mammalian cochlea." J. Physiol. 284, 261-271.

Sachs, M. B., and Abbas, P. J. (1974). "Rate versus level functions for auditory-nerve fibers in cats: Tone-burst stimuli." J. Acoust. Soc. Am. 56, 1835-1847.

Small, A. M. (1959). "Pure tone masking." J. Acoust. Soc. Am. 31, 1619-1625.

Small, A. M., and Busse, C. (1980). "Psychophysical tuning curves in forward masking." J. Acoust. Soc. Am. Suppl. 1 67, S9.

Stelmachowicz, P. G., and Jesteadt, W. (1984). "Psychophysical tuning curves in normal-hearing listeners: Test reliablity and probe level effects." J. Speech Hear. Res. 27, 396-402.

Terry, M., and Moore, B. C. J. (1977). "Suppression effects in forward masking." J. Acoust. Soc. Am. 62, 781-784.

Vogten, L. L. M. (1978a). "Simultaneous pure-tone masking. The dependence of masking asymmetries on intensity." J. Acoust. Soc. Am. 63, 1509-1527.

Vogten, L. L. M. (1978b). "Low-level pure-tone masking: A comparison of 'tuning curves' obtained with simultaneous and forward masking." J. Acoust. Soc. Am. 63, 1520-1529.

Weber, D. L. (1977). "Growth of masking and the auditory filter." J. Acoust. Soc. Am. 62, 424-429.

Weber, D. L. (1983). "Do off-frequency maskers suppress the signal?." J. Acoust. Soc. Am. 73, 887-893.

Weber, D. L., and Lutfi, R. A. (1986). "The unimportance of suppression." in Auditory Frequency Selectivity, edited by B. C. J. Moore and R. D. Patterson (Plenum, New York).

Widin, G. P., and Viemeister, N. F. (1979a). "Intensive and temporal effects in pure-tone forward masking." J. Acoust. Soc. Am. 66, 388-395.

Widin, G. P., and Viemeister, N. F. (1979b). "Short-term spectral effects in pure-tone forward masking." J. Acoust. Soc. Am. 66, 396-399.

Wightman, F. L., McGee, T., and Kramer, M. (1977). "Factors influencing frequency selectivity in normal and hearing-impaired listeners." in Psychophysics and Physiology of Hearing, edited by E. F. Evans and J. P. Wilson (Academic, London).

Zwicker, E., and Jaroszewski, A. (1982). "Inverse frequency dependence of simultaneous tone-on-tone masking patterns at low levels." J. Acoust. Soc. Am. 71, 1508-1512.

Zwislocki, J., Pirodda, E., and Rubin, H. (1959). "On some post-stimulatory effects at the threshold of audibility." J. Acoust. Soc. Am. 31, 9-14.

Complex interactions between pairs of forward maskers

Robert A. Lutfi
Waisman Center
University of Wisconsin
Madison, Wisconsin    53705

## ABSTRACT
The present study was conducted to determine to what extent the combined effect of two forward maskers can be predicted from addition of their individual effects.   The maskers were 50-Hz wide noise bands with center frequencies ranging from 1.8 to 2.2 kHz.   The signal was a brief, 2.0-kHz tone burst.   When the maskers were gated on and off together, the combination produced sometimes more and sometimes less masking than predicted depending on the particular pair and the relative amounts of masking produced by the individual members of the pair.   The greatest discrepancy occured, however, when the masker pair was presented simultaneously with the signal or when the  forward maskers were presented in sequence.   In the latter case, the obtained threshold exceeded the predicted threshold by as much as 34 dB.

Keywords: Forward masking, additivity

## INTRODUCTION
Over the last several years, there has been renewed interest  in masking by pairs  of  'equated'  maskers; maskers which when  presented separately  produce equal  amounts  of  masking  of  a  common signal.  Early studies revealed that the masking produced  by such  pairs  often  exceeds  the  simple  sum  of the masking  produced by the individual  members  of  the  pair  (Bilger, 1959; Green, 1967; Pollack, 1964).   In the study, by Pollack, this excess masking (beyond that predicted  by simple summation) amounted to as much as 19 dB.  The more recent studies reveal the  excess  masking  effect  to  be wide spread.  Large amounts  of excess  masking  have  now  been obtained for  various pairs  of  sequential forward  maskers  (Hanna  et al., 1982; Penner, 1980; Penner and  Shiffrin,  1980; Widin and Viemeister, 1980), forward and simultaneous  maskers  (Jesteadt et al., 1982; Jesteadt  and  Wilke, 1982), forward  and  backward  maskers  (Patterson, 1971;  Penner,  1980;  Robinson  and  Pollack, 1973; Wilson and Carhart,  1971), and  pairs  of simultaneous maskers (Canahl, 1971; Lutfi,  1983;  Moore, 1985; Nelson, 1979; Patterson  and Nimmo-Smith, 1980, Zwicker, 1954).   Indeed, only pairs of concurrent, forwards maskers have so far been found not  to  produce any excess masking (Jesteadt and Wilke, 1982; Neff and  Jesteadt,  1983).   Such   results challenge the traditional view of masking  which  assumes   that the effects of maskers are additive (Fletcher,  1940; Patterson and Green, 1978).   They have lead to  a  different  class  of  models in which the effects of maskers  undergo  a

compressive nonlinear transformation prior to summation (Penner, 1980; Penner and Shiffrin, 1980; Lutfi, 1983, 1935).

A strong assumption of these nonlinear models is that the amount of excess masking depends only on the relative effects of maskers in the pair, not on their particular temporal or spectral configuration (1). The assumption reflects the early stage of development cf these models. For instance, a basic difference in excess masking does appear to exist for simultaneous versus nonsimultaneous masker pairs (cf. Penner, 1980; Lutfi, 1983). There is also the issue as to why all combinations of equated maskers appear to produce excess masking except concurrent forward maskers. This latter issue seems the more critical since it would be unparsimonious to assert that auditory transduction is nonlinear except in this one case.

In light of the results for concurrent forward maskers, it seemed appropriate to explore further the effects of these maskers. We began by focusing on combinations of maskers with frequencies disparate from the signal frequency. In the studies by Jesteadt, Neff and Wilke, the effective masker components were always at or very nearly equal to the signal frequency. We also examined the effects of varying the relative amounts of masking produced by the individual maskers in the pair. The pattern of results to emerge from these experiments was complex. Various maskers combined to produce significant amounts of excess masking, no excess masking, or even a release from masking. We next examined the masking produced by a pair of simultaneous maskers, a pair of sequential forward maskers, and a simultaneous-plus-forward masker. The first two pairs produced the largest amounts of excess masking observed in this study, as much as 34 dB for the pair of sequential forward maskers. In contrast, the effects of the simultaneous and forward masker when paired were additive. The data are enough to dishearten those who would propose such elegant models as offered by Penner (1980), Penner and Shiffrin (1980), and Lutfi (1983).

## I. METHOD
### A. Stimuli
The signal in all conditions was a 10-ms, 2.0-kHz sinusoid, gated on and off with 5-ms Kaiser ramps. This ramp produces spectral sidelobes that are more than 70 dB down from the primary lobe within 20% of the primary lobe center frequency (see Childers and Duriing, 1975). The maskers were 200-ms, 50-Hz wide noise bands with variable center frequencies, they were also gated on and off with 5-ms Kaiser ramps. The long-term, power-spectra of the noise bands had skirts that fell over 1000 dB/octave near the passband. Three pairs of maskers were used. The center frequencies of the maskers in each pair were 1.8 and 1.9 kHz, 1.9 and 2.1 kHz, and 2.1 and 2.2 kHz. The maskers were gated on and off together with the onset of the signal following after a 5-ms silent interval. All stimuli were generated digitally and output through 14-bit DACs. Each masker was randomly sampled from a 3-s file on each presentation. The signal and each masker in the pair were played over separate DACs (10-kHz rate for each DAC). When only one masker was presented, 0s (corresponding to 0 voltage) were output through the DAC otherwise occupied by the second masker. The output of each DAC was low-pass filtered at 4.0 kHz, 120 dB/octave for each masker and 96 dB/octave for the signal. After mixing, the stimuli were amplified and were presented over TDH-49 headphones to the right ear of subjects seated in a double-wall, IAC sound-attenuated chamber.

## B. Procedure

Signal thresholds were obtained using a standard, two- interval, forced-choice adaptive procedure (Levitt, 1971). Each trial block began with the signal about 15 dB above masked threshold. Signal level decreased after two consecutive correct responses, and it increased for each incorrect response. The initial value of the step size was 6 dB; after two reversals in the direction of attenuation, the step size was changed to 2 dB. The trial sequence ended for each subject individually after a total of 18 reversals in signal attenuation had been recorded. The first two reversals were ignored and the levels of the remaining reversals were averaged to obtain a threshold estimate.

The relative amount of masking produced by the individual maskers in each pair was varied by varying the relative levels of the maskers in the pair. One masker was always fixed at 65 dB SPL, while the other varied from 30 to 80 dB SPL in 10 dB steps. An entire masking function (signal threshold versus masker level) was obtained for a single masker configuration before proceeding to the next. Typically, three masking functions were obtained within each daily 2-h session. After a single threshold estimate had been obtained for all conditions of the experiment a replication was performed.

## C. Subjects

Four university students were paid at an hourly rate to participate as listeners in the study. The ages of the subjects ranged between 18 and 23 years. All reported normal hearing and all had at least 10-h previous experience with the adaptive, two-interval, forced-choice task.

## II. RESULTS

The pattern of results was the same for all subjects, therefore the threshold estimates for each condition were averaged across subjects and replication. Fig. 1 gives the masking functions for all three masker pairs. In each panel, unfilled circles represent the masking function for the variable-level masker presented alone and filled triangles represent the masking function for the variable-level masker in the presence of the fixed-level masker (vertical bars represent one standard error on either side of the mean). The masking produced by the fixed-level masker alone is designated by the dashed line in each panel. The masking functions for the single variable-level maskers are quite typical of those obtained in the past (Egan and Hake, 1950; Vogten, 1978). They have a slope of 1 or slightly greater for maskers below the signal frequency, and a slope slightly less than one for maskers above the signal frequency.

Assuming simple summation of masking, the amount of masking produced by the masker pairs should never exceed by more than 3 dB the amount of masking produced by the more effective member of the pair. Also, masking by the pair should never fall below that of the more effective member. Exceptions to both of these predictions are evident in Fig. 1. For example, when the 1.9 and 2.1-kHz maskers separately produce equivalent amounts of masking (where the circles and dashed lines intersect), the amount of masking produced by the pair is about 10 dB greater than either masker alone. This is equivalent to 7 dB of excess masking. Excess masking is evident for the other masker pairs as well, although the amount of excess masking is never as large. For all masker pairs, there also appears to be a release from masking; the pair actually produces less masking than the more effective member of the pair. The release from masking is evident in the left-hand portion of each panel where the masking function for the pair dips

below that produced by the fixed-level masker (dashed line). In most cases, the release from masking amounts to only a few dB, however, at least in one case (fixed 2.2-, variable 2.1-kHz masker) it amounts to over 10 dB.

The overall pattern of results is made more clear in Fig. 2. The dependent variable in Fig. 2 is the difference between the amount of masking produced by the pair and the amount of masking produced by the more effective member of the pair. This relative amount of masking is plotted as a function of the difference in the amount of masking produced by each masker individually. The unfilled circles are from the condition wherein the lower frequency masker was fixed in level. The filled squares are from the condition wherein the higher frequency masker was fixed in level. The solid line in the middle of each panel gives the prediction based on simple additivity of masking. From Fig. 2 it is possible to see that excess masking results whenever the individual maskers in the pair produce roughly equivalent amounts of masking. The effect is largest for the 1.9+2.1-kHz pair and is small or nonexistent for the 1.8+1.9-kHz pair. In contrast to the excess masking, a release from masking results whenever the individual maskers in the pair produce largely discrepant amounts of masking; again, the largest release from masking being obtained with the 1.2+2.2-kHz pair.

Fig. 3 shows a similar set of data from a control condition in simultaneous masking. In this condition, the maskers were identical to the 1.8 and 1.9-kHz maskers used before, however, the signal was a 2.0-kHz sinusoid which was gated on and off with the maskers using 5-ms Kaiser ramps as before. These data are consistent with the data from numerous other studies which have obtained 10 dB or more excess masking for the combination of two simultaneous maskers (Canahl, 1971; Green, 1967; Lutfi, 1983; Nelson, 1979; Patterson and Nimmo-Smith, 1980).

III. DISCUSSION
A. Is forward masking additive?
The studies of Jesteadt and Wilke (1982) and Neff and Jesteadt (1983) have suggested that the effects of concurrent forward maskers are additive. This is an important result because many other combinations of maskers have so far been shown to produce large nonadditive effects in the form of excess masking. The present data show that some combinations of concurrent forward maskers can produce nonadditive effects, both in the form of excess masking and as a release from masking. In this section, we consider possible explanations for the apparent discrepancy between the present and past results.

Consider first the excess masking. In the present study, significant amounts of excess forward masking were obtained only for the 1.9+2.1 and 2.1+2.2-kHz masker pairs, and then only when the individual maskers in the pair produced nearly equivalent amounts of masking. We believe that a different factor is responsible for the excess masking in each case. For the 1.9+2.1-kHz pair, a likely cause is 'off-frequency listening' (Patterson and Nimmo-Smith, 1980). Because the ear's frequency resolving power is finite, there is some spread of excitation across auditory frequency channels. Consequently, the channels providing the best signal-to-noise ratio may sometimes be located off the frequency of the signal, away from that of the masker. This form of off-frequency listening is restricted whenever the frequencies of the maskers bracket the signal, as is true for the 1.9+2.1-kHz pair. The result is that the effect of such maskers in combination may exceed the simple sum of their effects in isolation. A different interpretation is required for the 2.1+2.2-kHz pair since these maskers do not bracket the signal. In this case, the excess masking could have

resulted from masking by an aural combination band generated at the signal frequency. This combination band would have been expected to produce significant amounts of masking only when the level of the 2.2-kHz masker was equal to or slightly below the level of the 2.1 kHz masker (Goldstein, 1967; Greenwood, 1971). As it happens, these are exactly the circumstances under which the excess masking for this pair is observed.

Consider next the release from masking. At first, one might be inclined to attribute this effect to suppression of the less intense masker by the more intense masker (Houtgast, 1974). However, there are two reasons why this interpretation is inadequate. First, for at least two of the masker pairs, the masker frequencies seem too close to yield measureable suppression effects (see Shannon, 1976.). Second, even if the less intense masker were completely suppressed, we should not expect the amount of masking to be any less than that of the more intense masker. In other words, there should be no release from masking except perhaps in the very rare instance when the less intense masker produced the greater amount of masking. A more likely interpretation is that the release from masking reflects the use of a 'quality difference cue' between the signal and the masker pair (Moore, 1980; Fastl and Bechly, 1981). Moore and others have presented evidence that a brief signal may be confused as part of the forward masker when, as in the present experiment, the signal and masker share a similar 'pitch-like' quality. Adding a second masker, which itself produces relatively lit le masking, gives the overall masker a more 'noise-like' quality, thereby lessening the chance of such confusions.

If one accepts the possibility of such confounding influences, then for only one condition of the present study can the results be safely compared to those of Jesteadt and Wilke (1982), and Neff and Jesteadt (1983). This would be the 1.8+1.9-kHz condition in which the individual maskers produce nearly equivalent amounts of masking. For this condition, the effects of the maskers do appear to be additive; if not, the discrepancy is very small. Thus, the data for this condition, at least, appear to be consistent with the previous data using concurrent forward maskers.

B. Excess masking as a failure of waveshape analysis.

Barring any confounding interactions between maskers, why is it that only the effects of concurrent forward maskers appear to be additive? One explanation may be made in terms of waveshape analysis. First, consider what happens when two nonconcurrent forward maskers are combined. When either masker is presented alone, the signal will be detected as a brief pertubation in waveshape at the end of the masker. Adding a second masker, which is separated from the first in time, will produce a second perturbation in close temporal proximity to that produced by the signal. This second perturbation may 'mask', be confused with, or otherwise interfere with that produced by the signal which in turn may result in excess masking. The situation is different for concurren forward maskers. Gating the forward maskers on and off together simply eliminates this second perturbation, and thus, the proposed cause of the excess masking. It is of interest to note that a similar type of waveshape analysis has been suggested as the cause of excess simultaneous masking (Lutfi, 1986).

The foregoing analysis is easily tested. If waveshape analysis is responsible for the excess masking obtained with nonconcurrent maskers, then it should be possible to both minimize and maximize the excess masking by selecting masker combinations which respectively minimize and maximize the difficulty of waveshape analysis. An additional experiment was conducted. The signal was a 20-ms, 2.0-kHz sinusoidal

burst. The nonconcurrent masker pair was a 200-ms noise band immediately followed by a 20-ms noise band (i.e. the offset of the first noise band corresponded to the onset of the second noise band at the zero voltage points). Both noise bands were 50 Hz wide and both were centered at 2.0-kHz. All stimuli were gated on and off as before with 5-ms Kaiser ramps. The levels of the maskers were also selected so that each individually produced the same amount of masking. The first panel of Fig. 4 shows the thresholds for two subjects when the signal was presented immediately following the 20-ms masker. The dashed line gives the prediction assuming additivity of masking. Note from the insert that the masker and signal+masker waveshapes are clearly discriminable when either masker is presented alone. When the two maskers are combined, however, the difference in waveshapes is much less clear and the signal may be perceived as a continuation of the masker. As much as 34 dB of excess masking is obtained in this condition, a record amount for the combination of these types of maskers. Perhaps, more interesting is that when the signal is presented simultaneously with the 20-ms masker (second panel), the excess masking is essentially eliminated. In this case, whether the maskers are presented individually or together, the masker and signal+masker waveshapes are never particularly easy to discriminate.

## C. Summary

Recent models have been successful in describing the masking produced by various masker pairs based only on the relative amount of masking produced by each member of the pair (Jesteadt, 1983; Lutfi, 1983, 1985; Penner, 1980; Penner and Shiffrin, 1980). The present study reveals several exceptions to the predictions of these models. The masking produced by pairs of concurrent forward maskers is found to depend not only on the relative amounts of masking, but, also on the particular combination of masker frequencies chosen. The pattern of results is complex suggesting that several different processes may have been involved. Among these are off-frequency listening, cueing effects, and masking by aural distortion products. Even when such factors can be ruled out, there is still found to be a large discrepancy in the amount of combined masking produced by equated pairs of concurrent versus sequential forward maskers. It may be that many of the results obtained using combinations of forward maskers can be accounted for in terms of differences among stimulus waveshapes.

## FOOTNOTE
(1) This assumption is contigent upon the condition that the effective components of the maskers are separated either in frequency or in time.

REFERENCES

Bilger, R.C. (1959). "Additivity of different types of masking," J. Acoust. Soc. Am. 31, 1107-1109.

Canahl, J.A. (1971). "Two- versus four-tone masking at 1000 Hz," J. Acoust. Soc. Am. 50, 471-474.

Childers, D. and Durling, A. (1975). Digital Filtering and Signal Processing (West, St. Paul) pp. 437-439.

Egan, J.P. and Hake, H.W. (1950). "On the masking pattern of a simple auditory stimulus," J. Acoust. Soc. Am. 22, 622-630.

Fastl, H. and Bechly, M. (1981). "Post masking with two maskers: Effects of bandwidth," J. Acoust. Soc. Am. 69, 1753-1757.

Fletcher, H. (1940). "Auditory patterns," Review of Modern Physics 12, 47-65.

Goldstein, J.L. (1967). "Audtitory nonlinearity," J. Acoust. Soc. Am. 41, 676-689.

Green, D.M. (1967). "Additivity of masking," J. Acoust. Soc. Am. 41, 1517-1525.

Greenwood, D.D. (1971). "Aural combination tones and auditory masking," J. Acoust. Soc. Am. 50, 502-543.

Hanna, T.E., Robinson, D.E., Shiffrin, R.M., and Gilkey, R.H. (1982). "Forward masking of diotic and dichotic clicks by noise," J. Acoust. Soc. Am. 72, 1171-1177.

Houtgast (1974). "Lateral suppression in hearing: A psychophysical study of the ear's capability to preserve and enhance spectral contrasts," Doctoral dissertation, Academishe Pers. B.V., Amsterdam.

Jesteadt, W. (1983). "A model of the combined effects of forward and simultaneous masking," J. Acoust. Soc. Am. 73, S91.

Jesteadt, W., Weber, D.L. and Wilke, S. (1982). "Interaction of simultaneous and forward masking for maskers and signals differing in frequency," J. Acoust. Soc. Am. Suppl. 71, S72.

Jesteadt, W., Wilke, S. (1982). "Interaction of simultaneous and forward masking," J. Acoust. Soc. Am. 71, S72 (A).

Levitt, H. (1971). "Transformed up-down methods in psychoacoustics," J. Acoust. Soc. Am. 49, 467-477.

Lutfi, R.A. (1983). "Additivity of simultaneous masking," J. Acoust.Soc. Am. 73, 262-267.

Lutfi, R.A. (1985). "A power-law transformation predicting masking by sounds with complex spectra," J. Acoust. Soc. Am. 77, 2128-2135.

Lutfi, R.A. (1986). "Two- versus four-tone masking, revisited," J. Acoust. Soc. Am., 80 422-428.

Moore, B.C.J. (1980). "Detection cues in forward masking," in Psychophysical, Physiological and Behaviora Studies in Hearing, edited by G. van den Brink and F.A. Bilsen (Delft U.P., Delft).

Moore, B.C.J. (1985). "Additivity of simultaneous masking, revisited," J. Acoust. Soc. Am. 78, 488-494.

Neff, D.L. and Jesteadt, W. (1983). "Additivity of forward masking," J. Acoust. Soc. Am. 74, 1695-1701.

Nelson, D.A. (1979). "Two-tone masking and auditory critical bandwidths," Audiology 18, 279-306.

Patterson, J.H. (1971). "Additivity of forward and backward masking as a function of signal frequency," J. Acoust. Soc. Am. 50, 1123-1125.

Patterson, R.D. and Green, D.M. (1978). "Auditory masking," in Handbook of Perception, Vol. VI, edited by E.C. Carterette and M.P. Friedman (Academic Press, New York).

Patterson, R.D. and Nimmo-Smith, I. (1980). "Off-frequency listening and auditory-filter asymmetry," J. Acoust. Soc. Am. 67, 229-245.

Penner, M.J. (1980). "The coding of intensity and the interaction of forward and backward masking," J. Acoust. Soc. Am. 67, 608-616.

Penner, M.J. and Shiffrin, R.M. (1980). "Nonlinearities in the coding of intensity within the context of a temporal summation model," J. Acoust. Soc. Am. 67, 617-627.

Pollack, I. (1964). "Interaction of forward and backward masking," J. Aud. Res. 4, 63-67.

Robinson, C.E. and Pollack, I. (1973). "Interaction between forward and backward masking: A measure of the integration period of the auditory system," J. Acoust. Soc. Am. 53, 1313-1316.

Shannon, R.V. (1976). "Two-tone unmasking and suppression in a forward-masking situation," J. Acoust. Soc. Am. 59, 1460-1470.

Vogten, L.L.M. (1978). "Simultaneous pure-tone masking: The dependence of masking asymmetries on intensity," J. Acoust. Soc. Am. 63, 1509-1527.

Widin, G.P. and Viemeister, N.F. (1980). "Masker interaction in pure-tone forward masking," J. Acoust. Soc. Am. 68, 475-479. -4

Wilson, R.H. and Carhart, R. (1971). "Forward and backward masking: Interactions and additivity," J. Acoust. Soc. Am. 49, 1254-1263.

Zwicker, E. (1954). "Die Verdeckung von Schmallbandgerauschen durch Sinustone,"
Acustica 4, 415-420.

FIGURE CAPTIONS

Fig. 1   Masking functions for the three pairs of concurrent forward maskers.  Unfilled circles give mean thresholds (4 subjects) in the presence of the variable-level masker alone.   Vertical bars represent one standard error on either side of the mean.   The dashed line denotes the threshold in the presence of the fixed-level (65 dB SPL) masker alone.   Filled triangles give thresholds when the variable- and fixed-level maskers are combined.

Fig. 2   The ordinate gives, for each masker pair, the amount of masking above or below that produced by the most effective member in the pair.   The abscissa gives the difference in the amount of masking produced by the individual members of the pair. Unfilled circles denote that the fixed-level masker was the lower frequency masker.   Filled squares denote that the fixed-level masker was the higher-frequency masker.   The solid line in the center of each panel gives the prediction based on simple additivity of masking.

Fig. 3   Same as Fig. 2, except the two maskers were presented simultaneously with the signal.

Fig. 4   Panel A: Signal threshold in the presence of a pair of sequential forward maskers is plotted as a function of the threshold in the presence of either masker alone (different symbols represent data from two subjects).   The dashed line gives the prediction based on additivity of masking.   Panel B: Same as panel A, except the signal was presented simultaneously with the trialing masker.   The insert in each panel shows the waveshapes of the signal and both maskers combined.   The cross-hatched region denotes the signal.
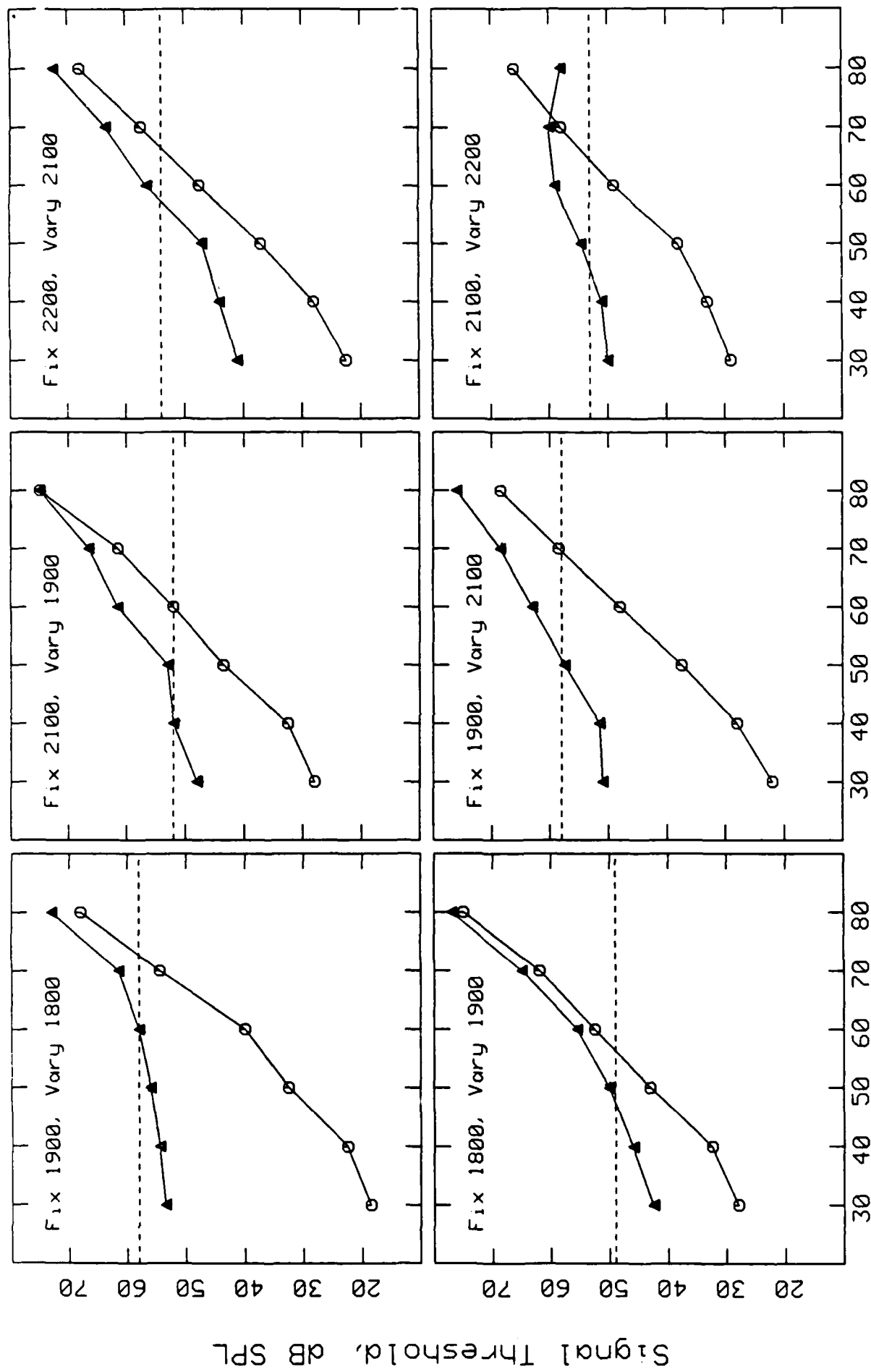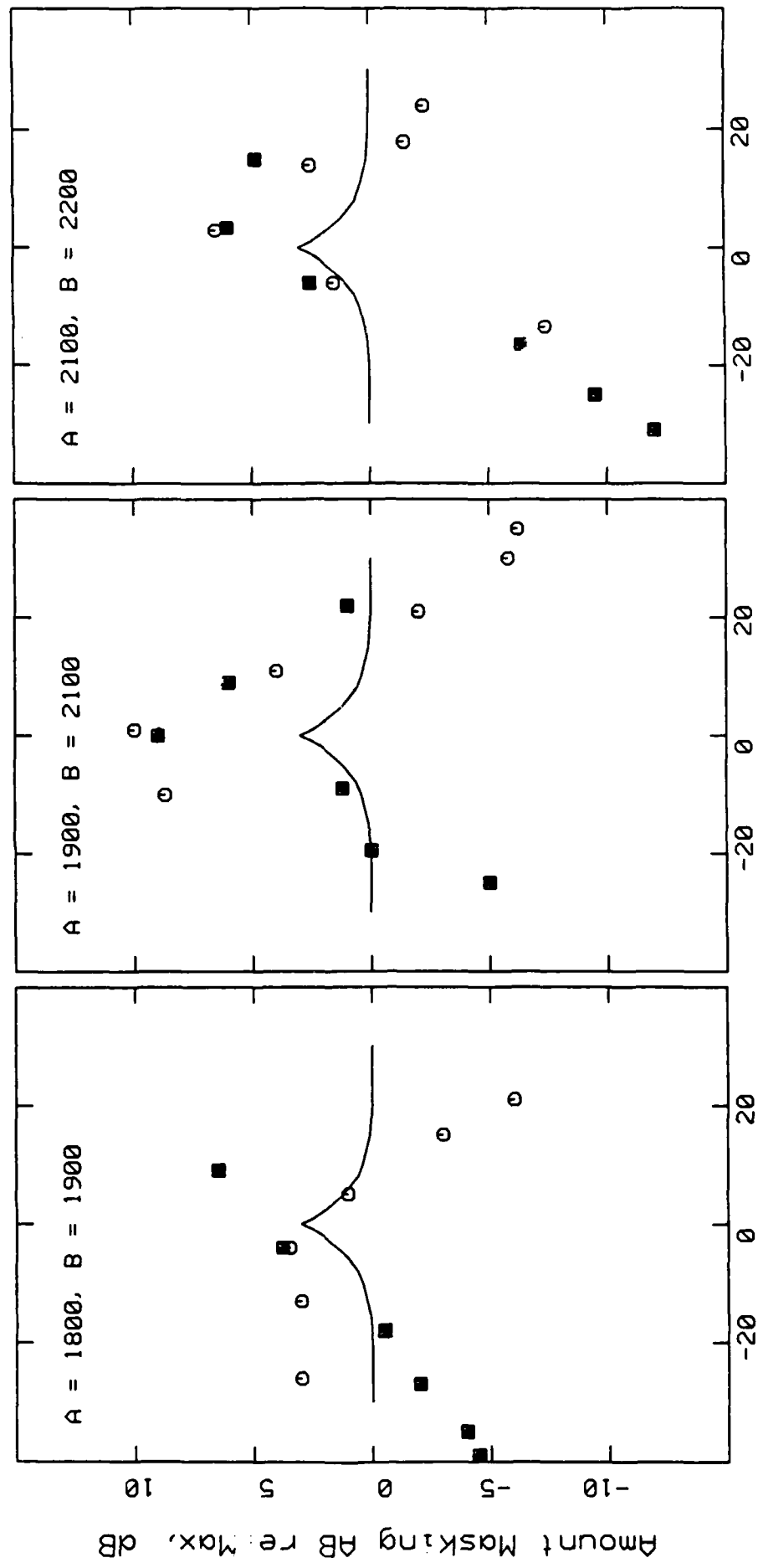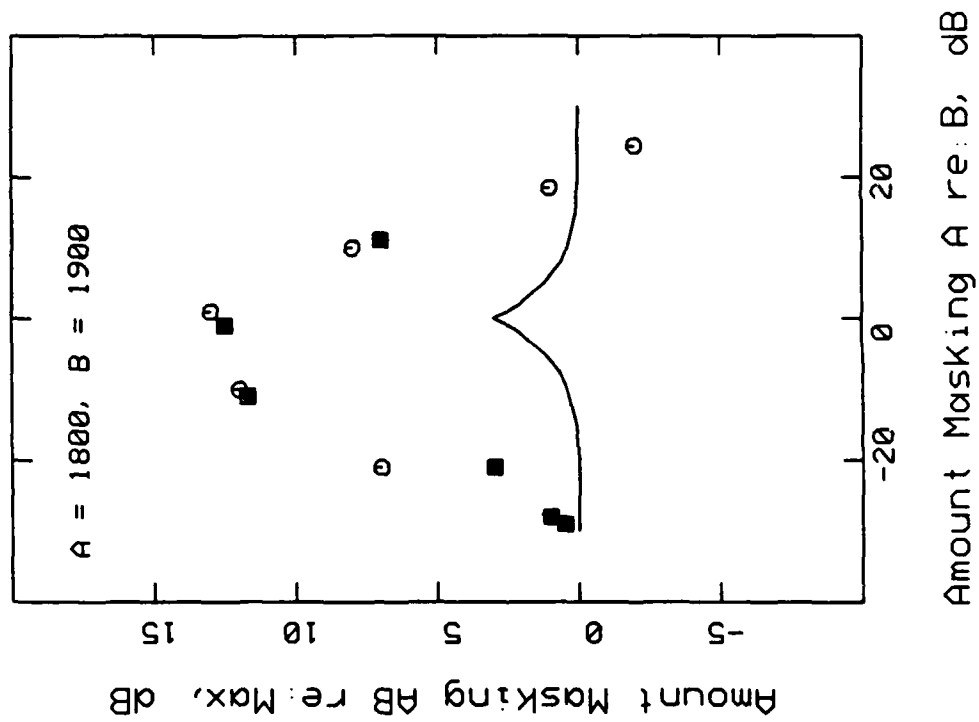
FIG 1

Masker Level, dB SPL

Signal Threshold, dB SPL

Fix 1900, Vary 1800

Fix 2100, Vary 1900

Fix 2200, Vary 2100

Fix 1800, Vary 1900

Fix 1900, Vary 2100

Fix 2100, Vary 2200

FIG. 3

A = 1800, B = 1900

Amount Masking A re: B, dB

Amount Masking AB re:Max, dB

FIG 3

Fig 4

Masked Threshold A, dB SPL

Masked Threshold AB, dB SPL